MEASUREMENT OF MOLECULAR CONFORMATIONS AND DYNAMICS USING SINGLE MOLECULE FLUORESCENCE TECHNIQUES

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Huimin Chen

January 2012

MEASUREMENT OF MOLECULAR CONFORMATIONS AND DYNAMICS USING SINGLE MOLECULE FLUORESCENCE TECHNIQUES

Huimin Chen, Ph.D.

Cornell University 2012

The use of fluorescence spectroscopy to study biological problems has gained popularity over the past few decades. Beyond a spatial understanding provided by microscopy, fluorescence techniques like Fluorescence Correlation Spectroscopy (FCS) and fluorescence lifetime spectroscopy can also elucidate the important temporal dynamics of molecular conformations.

We have applied FCS to study the conformational fluctuations in a model protein apomyoglobin. By pushing the technical limitations of FCS, we were able to observe conformational dynamics spanning two orders of magnitude in time (10^{-6} to 10^{-3} seconds). We found that the amplitude of fluctuations changes as the molecule becomes unfolded, with principal shifts of amplitudes and timescales occurring at the transition across the molten globule state. We also measured the diffusion of apomyoglobin, confirming theoretical predictions of less compaction of the molecule upon acid denaturation.

We were able to observe the fluctuations in apomyoglobin using FCS due to the quenching of an N-terminal labeled Alexa488 fluorophore by contact with various amino acids. We showed that quenching can occur with up to four amino acids. We investigated the mechanisms of quenching using a combination of fluorescence intensity and lifetime measurements. We showed that quenching takes place through a combination of static and dynamic mechanisms. Our results demonstrate that care needs to be taken when making quan-

titative measurements of fluorescently labeled proteins.

Coupling single molecule functionality to fluorescence techniques allows researchers to discern subpopulations within ensembles. We used single molecule Förster Resonance Energy Transfer (smFRET) to measure the end-to-end distances in single stranded nucleic acids as a probe of the flexibility. We also measured the radius of gyration using small angle X-ray scattering, and extracted parameters of polymer properties that fit the data from both techniques. We observed clear differences between our model single stranded DNA (ssDNA) and RNA (ssRNA). We also observed a difference between the screening efficiency of monovalent and divalent cations. By characterizing the intrinsic differences in nucleic acids and its dependence on the ionic strength, we hope to improve our understanding of the mechanisms of RNA folding and the role of ions in the process.

BIOGRAPHICAL SKETCH

Huimin Chen was born and raised in Singapore, where she spent her childhood battling the academic rigors of the Singapore education system. She had no idea what she wanted to do in life, but went along with the flow when all her classmates applied to universities overseas. After whittling down her choices based on photographs of the different college on their brochures, she enrolled at Cornell University in the fall of 2000. At first she thought she would study an engineering discipline and upon graduation, she would return home to a stable job with a middle class lifestyle in Singapore. However, a particularly inspiring professor in her freshman Physics class showed her for the first time in her life that science was fun and more interesting than just memorizing equations. One semester later, she became an Engineering Physics major. She decided late in her junior year that she wanted to go to graduate school, and so she sought out biophysics professors for some research experience. After 23 rejection letters, she serendipitously joined the laboratory of Watt W. Webb where she wrote her undergraduate honors thesis, and subsequently decided to stay on for a PhD after she obtained her bachelors degree in Engineering Physics in spring 2004. Throughout the years, Huimin has worked on variety of research projects ranging from UV crosslinking in nucleic acids to protein folding to neurodegenerative diseases. Upon graduation, Huimin hopes to continue her research as a postdoctoral scholar.

For my parents, who must have been wondering why I had to be so far away from home all these years...

ACKNOWLEDGEMENTS

I consider this chapter the most important and yet the most difficult part of the thesis to write. Its importance is obvious, because it takes an entire university to groom and train a graduate student and many thanks are due. However, it is difficult because how do you capture your complicated feelings of gratitude in only so many words? I will try my best here...

I am forever indebted to my advisor Watt W. Webb for taking a chance on me as an undergraduate student, and then for persuading me to stay at Cornell for my PhD. I thank him for also stocking the lab with a team of talented individuals from whom I have benefitted immensely. His broad scientific vision, uncanny ability to read people, his life stories, advice on work and life, and generous guidance have set a lofty standard to which I will aspire to be as an advisor and boss. I also thank him for understanding the importance of breadth in an education. As a result I was able to study the Japanese language throughout my PhD, an undertaking that brought me immeasurable joy and pride.

The 'characters' in the Webb group over the years have provided me with invaluable scientific advice and many memories. The spirit and camaraderie amongst the current members and alumni, young and old is rare and many thanks go to them. To Becky and Warren Zipfel, for always knowing the answers. To Xanthipe Jordanides, for showing me what a great scientist was. To Liz 'Ma'am' Rhoades, for being a great mentor and colleague, and for showing me that with an amazing work ethic, anything was possible. I hope I was a good student. To Mark Williams for all his administrative, formatting and fishing genius. To Tobias Baumgart, Mike Levene, Dan Dombeck, Dan Larson, Harsh Visrasrao, Iain Tullis, Jonas Korlach, Chris Fecko, Huizhong Xu, Jie, Jesse McMullen, Alex Kwan, Valerie Anderson, Elaine Farkas, Ina Pavlova, Chris Brown,

Dimitre Ouzounov, Avtar Singh and many others for the great company and memories. To Saad Ahsan, for putting up with me and providing me with a great teaching opportunity - I tried to be the best mentor I could bearing in mind all the great role models I had before me and I hope I was at least half as successful.

I am also grateful to Lois Pollack for actively pursuing collaborations with the Webb group and 'adopting' me into her wonderful group. Lois, together with Suzette Pabit, Steve Meisburger and Josh Blose, have taught me everything I know about nucleic acids. To Steve, you have given new meaning to 'enjoyable collaboration', and I learnt so much from working with you. The successful collaboration gave rise to Chapter 4, and an appreciation and budding interest in nucleic acids biology within me. I also thank Peng Chen for serving on my thesis committee.

Collaborations can be a rewarding part of academic research and I have directly benefitted greatly from the generosity of people around me. For the work done in Chapter 2, I thank Stewart Loh, Jim Butler and Margaret Stratton from Upstate Medical University of SUNY. You have all been very patient over the course of this collaboration that lasted over ten years, most of the time providing us with precious protein sample with very little results. I thank Mitk'El Santiago-Berrios and Hector Abruna from the Chemistry department for responding so enthusiastically to an email from a total stranger. Their efforts resulted in the beautiful electrochemical data described in Chapter 3. Over the years I have also been helped by John Grazul at the Cornell Center of Materials Research (CCMR) on the transmission electron microscopy, by Ellen Keene on circular dichroism and by Sned and Nate at the machine shop.

I am grateful to have friends and colleagues who cared and supported me

throughout my PhD. I thank my sister for reading my papers and listening to my talks over and over again, and in the process becoming half a biophysicist herself. I thank my parents for indulging me, and never questioning what I choose to do. Finally, to Shangyen, who sometimes knows me better than I know myself, thank you for the companionship and the patience.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

The study of molecular biology has been greatly enhanced by techniques like X-ray crystallography and nuclear magnetic resonance (NMR), which provide images of biomolecular structures like proteins and nucleic acids at the atomic level. Structural information can provide a wealth of information about modes of interaction between molecules, and the mechanisms through which they function. However, these traditional techniques are limited in two ways. Firstly, they do not provide much information about the temporal dynamics of the molecules. Secondly, these techniques tell us only about the ensemble averaged state, but give us little information about subpopulations and disordered states. Fluorescence studies have the ability to fill in some of these information gaps.

As the fluorescence properties of a fluorophore are extremely sensitive to its local environment, one can gather much information about the molecule's surroundings by looking at the fluorescence intensity in the time and spectral domain. In this dissertation, I show through the use of various techniques, that fluorescence studies can be used to (1) measure the timescales and amplitudes of structural fluctuations within the folded and unfolded states of a protein, (2) probe the contacts between two parts of a protein and (3) measure the chain flexibility of a molecule that does not form an ordered structure.

While static images of structure from X-ray crystallography or NMR give important information about the folded state, the reality is that molecular structures fluctuate and 'breathe'. These structural fluctuations are important in protein-protein interactions, as well as protein folding and unfolding. These

stochastic motions are believed to be key in funneling an unfolded protein towards its correctly folded form, and the timescales and modes of these motions can give insight into the mechanisms that underlie the high fidelity of protein folding. Knowledge of the folding mechanisms can then be used for rational design of therapeutic protein molecules. Furthermore, protein structure fluctuations are also crucial for understanding protein misfolding, a phenomenon that has been implicated in a host of human diseases like Parkinson's and Alzheimer's diseases. In these diseases, proteins inexplicably misfold and aggregate, forming amyloid structures which are hallmarks of theses diseases. These diseases are widely prevalent yet the triggers of aggregation are unknown. We are motivated to understand the mechanisms and triggers of misfolding through measuring the modes and timescales of molecular fluctuations.

In Chapter 2, we use a model protein apomyoglobin, which while not known naturally to undergo amyloid formation, has been observed to do so under aggressive conditions. We use fluorescence correlation spectroscopy (FCS) to measure the timescales of chain motions within the protein under native and unfolding conditions. To indirectly observe the fast conformational fluctuations, we exploit the contact quenching that occurs when the N-terminus labeled fluorophore comes in close proximity of certain amino acid residues within the same protein. The fast contact quenching allows us to measure timescales associated with chain motions with FCS.

The mechanism of quenching that enables us to probe conformational fluctuations was not well known or understood. In Chapter 3, we introduce the basic mechanisms of quenching when molecules approach each other closely.

We show that specific amino acids can have a large quenching effect on a commonly used fluorophore during collisional contact. We also elucidate some of the mechanisms involved in the quenching process through a combination of fluorescence lifetime and intensity measurements. The results of Chapter 3 provides a cautionary reminder that fluorophores can be altered by the molecules they are attached to and experiments should take these effects into consideration when quantitative measurements of fluorescence intensities are to be made.

Another popular fluorescence technique that can potentially measure the timescales of conformational fluctuations is Forster Resonance Energy Transfer (FRET). This technique that takes advantage of the exquisite sensitivity of fluorescence properties to measure the separation between two spectrally and spatially distinct fluorophores. When two fluorophores are close together, through-space energy transfer from the excited state of the donor fluorophore to the acceptor fluorophore can occur through dipole-dipole coupling. The excited donor fluorophore does not emit fluorescence but instead the acceptor fluorophore does so, with an efficiency that depends on the separation between them. This technique is extremely sensitive to separations in the range of 20-70 Å, making it an invaluable tool for probing conformational changes or displacement of molecules. In particular, the single molecule implementation of this technique has allowed researchers to distinguish subpopulations within the ensemble. This information was previously obscured in ensemble bulk measurements.

In Chapter 4, we utilize single molecule FRET (smFRET) to measure the separation between a Cy3 donor fluorophore attached to one end of a short single stranded nucleic acid and a Cy5 acceptor fluorophore attached to the other end.

This gives us an indirect probe of the flexibility of the nucleic acid chain. As it is challenging to ensure that a molecule is labeled with both the donor and the acceptor fluorophore needed for FRET, smFRET is needed to distinguish between donor-labeled-only, acceptor-labeled-only and double-labeled molecules. Using a combination of smFRET and small angle X-ray scattering (SAXS), we compared the chain flexibilities of single stranded DNA (ssDNA) and single stranded RNA (ssRNA), in the presence of monovalent and divalent counterions. We showed that the flexibility of nucleic acid chains is very sensitive to its local environment, and highlights the extra care that should be taken when experiments involve tethering molecules down to a surface.

Finally, Chapter 5 is an overall summary and describes future outlook of the projects described in this dissertation.

CHAPTER 2

DYNAMICS OF EQUILIBRIUM STRUCTURAL FLUCTUATIONS OF APOMYOGLOBIN MEASURED BY FLUORESCENCE CORRELATION SPECTROSCOPY*

2.1 Introduction

A major challenge of structural biology is elucidation of the complex pathways by which proteins attain their folded, functional structures, or in the case of many proteins, misfolded structures. Understanding requires not only characterization of the structures populated during folding, but also the dynamics of the polypeptide chain as its motions become restricted to progressively narrower ranges as folding reaches toward completion. Of particular interest are the intrachain contacts formed in unfolded or partially folded molten globule states, which may limit the conformational space sampled by the chain, directing it toward its final structure. However, because the relevant chain motions in structural fluctuations can span time scales from picoseconds to seconds, they may be elusive to traditional techniques (such as NMR) or lost in the dead time of stopped flow measurements, and thus have been difficult to characterize.

Extensive characterization of both the thermodynamics and kinetics of the sperm whale apomyoglobin (apoMb) structure by numerous techniques has made it a paradigm for protein folding studies [82]. ApoMb is generated by

*Originally published as: Chen, H, Rhoades, E., Butler, J. S., Loh, S. N., Webb, W. W. (2007). "Dynamics of equilibrium structural fluctuations of apomyoglobin measured by fluorescence correlation spectroscopy," Proceedings of the National Academy of Sciences of the United States of America1104(25), 10459-10464.

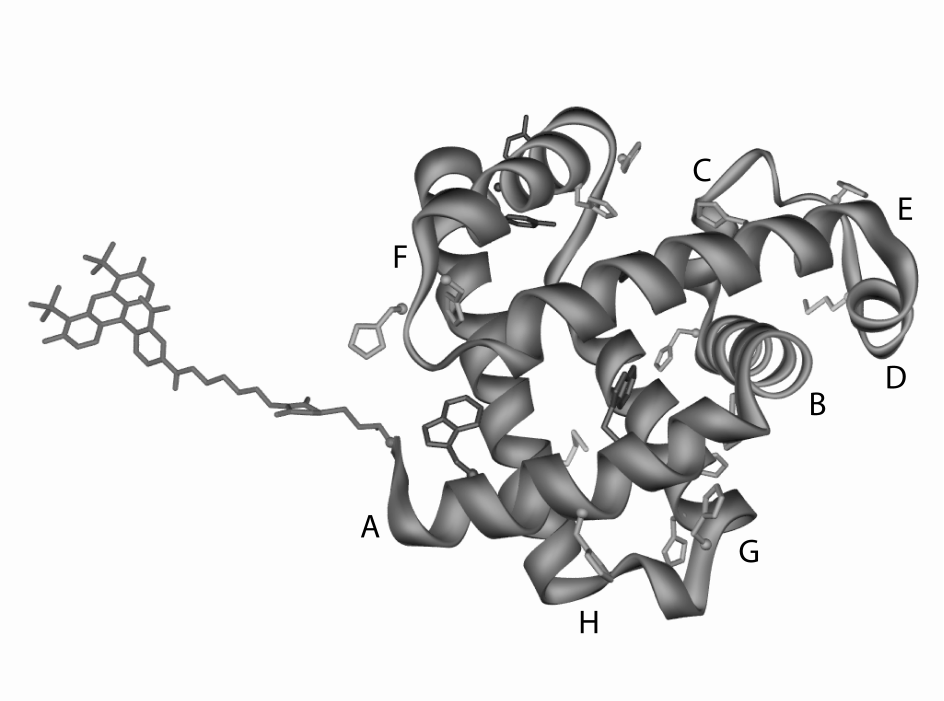

Figure 2.1: Ribbon representation of apoMb labeled with Alexa 488 (green) at the amino terminus, with helices labeled A-H. Potential side chain quenching groups are shown and colored as follows: Tyr (red), Trp (purple), Met (cyan), and His (blue).

removing the noncovalently bound heme group from myoglobin (Figure 2.1). ApoMb has provided a unique insight into the folding reaction because it undergoes a three-state pH-induced unfolding transition via stable molten globule intermediates.

It has been reported that apoMb can be induced to transition from its native all-helical structure to an aggregated β-sheet form that is characteristic of amy-

loid deposits, which are associated with a broad range of diseases [25, 12]. This observation, which motivates the research presented here, implies the potential for insight into protein misfolding by studying the structural fluctuations associated with folding pathways [25].

With the ability to access time scales from nanoseconds to seconds, fluorescence correlation spectroscopy (FCS) has emerged as a powerful technique for characterizing dynamics of molecules of biochemical and biophysical interest [62, 24, 61, 21, 92]. Although commonly used to measure the diffusion-associated properties of biomolecules, FCS can also characterize fluctuations in fluorescence intensity caused by chemical kinetics or photophysics [62, 24, 61, 21, 33, 85, 114]. A number of studies have demonstrated the utility of FCS in monitoring folding or function related conformational fluctuations in RNA [46, 47], DNA [58, 111, 20, 43, 42], polypeptides [77], and, of particular interest to this study, proteins [9, 8, 113, 93, 65]. Chattopadhyay et al. [9, 8] followed a pH-induced unfolding transition in a small fatty acid binding protein, finding that the decrease of the amplitude of a $\approx$35 μs fluctuation mirrored the unfolding of the protein and later demonstrating that their observed motions were diffusive in nature. Similarly, Rischel et al. [93] were able to put an upper bound of $\approx$4 μs on the time scale of fluctuations in denatured cytochrome c. Here, our use of FCS to measure equilibrium structural fluctuations of apoMb challenges current instrumental capabilities and tests the limits of FCS autocorrelation methods.

ApoMb can be denatured conveniently by the addition of acid. The native state (N state) of apoMb is formed above pH 6. Its structure is similar to that of the heme-bound protein except for enhanced flexibility and partial unfold-

ing of the F-helix [22]. At pH 4.2, apoMb exists as a mixture of two partially folded forms, I_a and I_b [41], which are classic molten globules in that they contain extensive secondary structure but loosely packed cores [44, 37]. Structure in I_a is limited to helices AGH, and to a lesser extent, helix B; I_b is similar except that helix B is more structured [31]. I_a and I_b equilibrate on the 1- to 10-ms time scale at pH 4.2; however, the equilibrium constant has not been measured [112], thus "I" refers to the equilibrium mixture of I_a and I_b. At pH 2.5, apoMb exists in an unfolded state (U state) that nonetheless contains some helical secondary structure [38, 117, 23].

In this study, we apply FCS to measure the equilibrium chain dynamics of apoMb as a function of pH. The V1C mutation at the amino terminus of the protein allows for specific labeling with the fluorophore Alexa 488 (apoMb V1C-AL488). The fluorescence emission of Alexa 488 is sensitive to specific interactions with quenching groups in the protein chain, thereby reporting on the contact dynamics of the N terminus of the protein. We find that the correlation curves for these effects can be best represented by two exponential relaxation times above pH≈4.1, whereas the low pH curves require three relaxation times, allowing us to resolve fluctuations on timescales ranging from ≈3 to ≈200 μs. Although the relaxation times of the motions are relatively insensitive to pH, the amplitude of the correlation signal at each pH can be attributed to increase in fluctuations as the protein unfolds. In addition, we extract the diffusion time of apoMb at each pH as a measure of the relative hydrodynamic compactness of the structures populated.

2.2 Materials and Methods

2.2.1 Protein expression and labeling

ApoMb-V1C was cloned, expressed, and purified as described [31]. In all subsequent experiments, the buffers used are 5 mM citrate, 20 mM NaCl, with pH adjusted from 2.6 to 6.3. ApoMb-V1C was reduced with tris(2-carboxyethyl)phosphine and labeled with Alexa 488 C5-maleimide (Invitrogen Corp., Carlsbad, CA) (apoMb-V1C-AL488) by incubating for 4 h in 4 M guanidine hydrochloride (GdnHCl, pH 7.0) at room temperature. Unconjugated dye was removed by desalting into 4 M GdnHCl (pH 7.0) using a PD10 column (Amersham, GE Healthcare Biosciences, Pittsburg, PA). The protein was then dialyzed extensively against double-distilled H_2O and lyophilized. Lyophilized protein was dissolved (final concentration of 20 μM) in 5 mM citrate (pH 2.6), 20 mM NaCl. This stock solution was aliquoted, flash frozen in liquid nitrogen, and stored at -20°C. Prior to measurements, an aliquot of protein was thawed at room temperature and dialyzed overnight at 4°C against 5 mM citrate (pH 2.6), 20 mM NaCl to remove any remaining unconjugated Alexa 488. The dialyzed protein was stored at 4°C and used within 10 days.

2.2.2 Steady state fluorescence and CD measurements

Fluorescence data were collected on a Fluoromax-3 fluorimeter (Horiba Jobin-Yvon, Edison, NJ) at 20°C. Alexa 488 was excited at 488 nm and emission spectra were recorded from 500 nm to 600 nm. At each pH value, fluorescence intensities were integrated and normalized to that of the native protein (pH 6.3).

The reported values were corrected for the absorption of Alexa 488 measured at each pH value. CD data were recorded on a Model 202 spectropolarimeter (Aviv Biomedical, Lakewood, NJ), at 20°C. Wavelength was set to 222 nm. For both fluorescence and CD experiments, WT and V1C concentrations were 5 μM, and V1C-AL488 concentration was 0.5 μM. CD experiments employed a 2 mm path length cuvette (WT and V1C) or a 1 cm cuvette (V1C-AL488). As a further test, we measured the stability of native apoMb-V1C-AL488 at pH 6.0 by urea denaturation experiments.

2.2.3 FCS measurements

A solid state laser with 488 nm emission (Coherent Inc., Santa Clara, CA) was directed into an IX71 microscope (Olympus America Inc., Center Valley, PA), reflected by a 510 nm long pass dichroic, and focused into the sample by a 1.15 NA 40x UApo/340 water objective. Fluorescence was collected through the objective and separated from the excitation light by the dichroic in conjugation with a 575±150 nm band pass filter. The emission light was focused by a 150 mm lens onto a 100 μm-diameter 50/50 beamsplitter optical fiber (OZ Optics, Ottawa, Ontario) which served as the confocal pinhole. The divided fluorescence emission was sent separately to two avalanche photodiodes (APDs) (SPCM-AQR, PerkinElmer, Wellesley, MA). Cross-correlations selected to reject double pulsing by the APDs were calculated from the digitized signals with a USB correlator (Correlator.com, Bridgewater, NJ).

The laser beam waist prior to entering the objective was 1 mm, such that the back aperture of the objective was underfilled, resulting in an enlarged and elon-

gated focal volume [36]. A concern in fitting the FCS data is that the enlarged focal volume could deviate significantly from the expected prolate Gaussian ellipsoid model approximation. This effect could introduce artifacts in the fitting dynamics. Therefore, to exclude this problem, a reference solution of Alexa 488 hydrazide dye was measured daily prior to any protein measurements to calibrate the instrument and to characterize the focal volume. These data were fit to a correlation function expected for one freely diffusing species with all the parameters allowed to float to determine ω, which describes the dimensions of the focal volume. ω was consistently found to be 7.7 and thus was fixed to this value for all subsequent fitting of protein data, while all other parameters were allowed to float. The Alexa 488 hydrazide data fit very well to the model of one freely diffusing species, with flat residuals down to a timescale of 100 ns (data not shown). The goodness of fit for the single diffusing species indicates a lack of artifacts due to the enlarged focal volume, enabling us to fit subsequent data that exhibit more complex dynamics with confidence.

Low illumination intensities were used to avoid triplet state photophysics. The amplitudes of the fast components, which are comparable in timescale to triplet state dynamics, did not increase even at powers significantly higher than those used in our experiments, indicating that they are due to protein dynamics, not Alexa 488 photophysics.

ApoMb-V1C-AL488 samples were prepared by dilution of 0.5-2.0 μl of the stock protein solution into 250 μl of the appropriate pH buffer for an estimated final protein concentration of $\approx$50 nM. Measurements were made in eight-chambered coverglass slides (LabTek, Nunc, Rochester, NY). Prior to use, the chambers were oxygen plasma-cleaned and then incubated with polylysine

conjugated polyethylene glycol (PLL-PEG) to minimize adhesion of the protein to the surfaces of the chamber [45]. The chambers were rinsed thoroughly with deionized H_2O prior to the addition of the protein solutions. Focal volumes did not overlap chamber walls. For a given pH solution, a typical data set consists of 100 correlation curves of 10 seconds each (for a total of 1000 seconds). When necessary, the curves were normalized to the intensity traces to account for any fluctuations in protein concentration between measurements, and then averaged to yield a single curve. The averaged curves were fit to the various models, described in the following section, using the inverse square of the standard deviation as a weighting factor. Data sets were collected on at least three different days from three different protein samples. All measurements were made at room temperature.

2.2.4 Fitting models

The autocorrelation function G(t) is defined as:

$$G(\tau) = \left\langle \frac{\delta F(t)\delta F(t+\tau)}{F(\tau)^2} \right\rangle \tag{2.1}$$

where F(τ) is the fluorescence measured at delay time τ after time t and δF(t)=F(t)-<F(t)>. For solutions that contain only a single species diffusing in three dimensions, the correlation curves can be fitted with:

$$G_D(\tau) = \frac{1}{N}\left[1 + \frac{\tau}{\tau_D}\right]^{-1}\left[1 + \frac{\tau}{\omega^2\tau_D}\right]^{-\frac{1}{2}} \tag{2.2}$$

where N is the mean number of molecules in the focal volume, τ_D is their diffusion time, and ω is the experimentally determined ratio of the axial to radial di-

mensions of the observation volume. If the species of interest is undergoing conformational fluctuations amongst multiple states that can be distinguished by their relative fluorescence quantum yields, and assuming no detectable change in τ_D between the two states, then the correlation function can be written as [61, 80, 70]:

$$G(\tau) = G_D(\tau)G_F(\tau) \tag{2.3}$$

where $G_F(\tau)$ represents the contribution of the fluctuations to the correlation function. In our optical set-up, the back aperture of the objective was underfilled to create an enlarged focal volume resulting in an increase in the diffusion time of the protein (τ_D). A longer τ_D allows for better separation of the contributions to the correlation function due to diffusion (G_D) from those due to the conformational fluctuations (G_F), thereby accommodating accurate determination of the timescales of the fluctuations. In its most general form [80, 115]:

$$G_F(\tau) = \sum A_i exp(\lambda_i \tau) \tag{2.4}$$

where λ_i and A_i are the eigenvalues and eigenvectors of the matrix that describes the time evolution due to kinetics of the species of interest.

In modeling the simplest case, a species fluctuating between two states, the equation used is generally in one of these two forms shown below:

$$\begin{aligned} G_F(\tau) &= 1 + A_{eq} exp(-\frac{\tau}{\tau_A}) \\ &= \frac{1 - A + A exp(-\frac{\tau}{\tau_A})}{1 - A} \end{aligned} \tag{2.5}$$

where τ_A is the relaxation time of the isomerization, A is the fraction of molecules in the isomerized state, and A_{eq}=A(1-A) can be explicitly written out as a function of the equilibrium rate constants and the various quantum yields of the states [61, 80].

For more complex systems, where there are multiple fluorescent states of essentially equal diffusivities with multiple characteristic timescales that describe the transitions between these states, Equation 2.4 can be expanded in a straightforward manner:

$$G(\tau) = G_D(\tau)\left(\frac{1 - \sum_i (A_i - A_i exp(-\tau/\tau_{A_i}))}{1 - \sum_i A_i}\right) \tag{2.6}$$

In addition to the simple diffusion model (1D), we also fit to a model with a variable number (i) of kinetic components (F) to account for the broad spectra of observed fluctuations (1DiF).

For two diffusing components (2D), we will fit with an equation defined by:

$$G(\tau) = B \times G_{D_1}(\tau) + (1 - B) \times G_{D_2}(\tau) \tag{2.7}$$

where B is the fraction of molecules with one diffusion time.

2.3 Results

2.3.1 Labeling does not perturb protein structure

To determine the extent to which the Alexa 488 label perturbs equilibrium folding of apoMb, we monitored acid-induced unfolding of WT and V1C-AL488 by using CD (Figure 2.2). ApoMb WT unfolds in the well documented three-state manner. The N/I transition occurs with a midpoint (pK_1) near pH 4.7. The characteristic molten globule state (I) is maximally populated at pH 4.1. I unfolds to U with a midpoint (pK_2) near pH 3.4. Introduction of the V1C mutation and the dye group does not change pK_1, pK_2, or the molar ellipticities of N, I, or U. This finding indicates that the Alexa 488 label does not perturb the structures or stabilities of the species populated throughout the acid-induced unfolding reaction.

We characterized the acid-induced unfolding reaction and the accompanying dependence of the fluorescence of apoMb-V1C-AL488 by 488 nm bulk fluorescence (Figure 2.3). The fluorescence intensity (corrected for the changes in Alexa 488 absorption caused by pH) decreases as the protein unfolds from N to I, and decreases further as I unfolds to U , indicating that at position 1, Alexa 488 serves as a good probe of apoMb denaturation. Unfolding curves of apoMb-V1C-AL488 monitored by Alexa 488 fluorescence (inset of Figure 2.3) and apoMb WT monitored by Trp fluorescence (data not shown) are nearly superimposable, confirming that the label does not change stability of the native state.

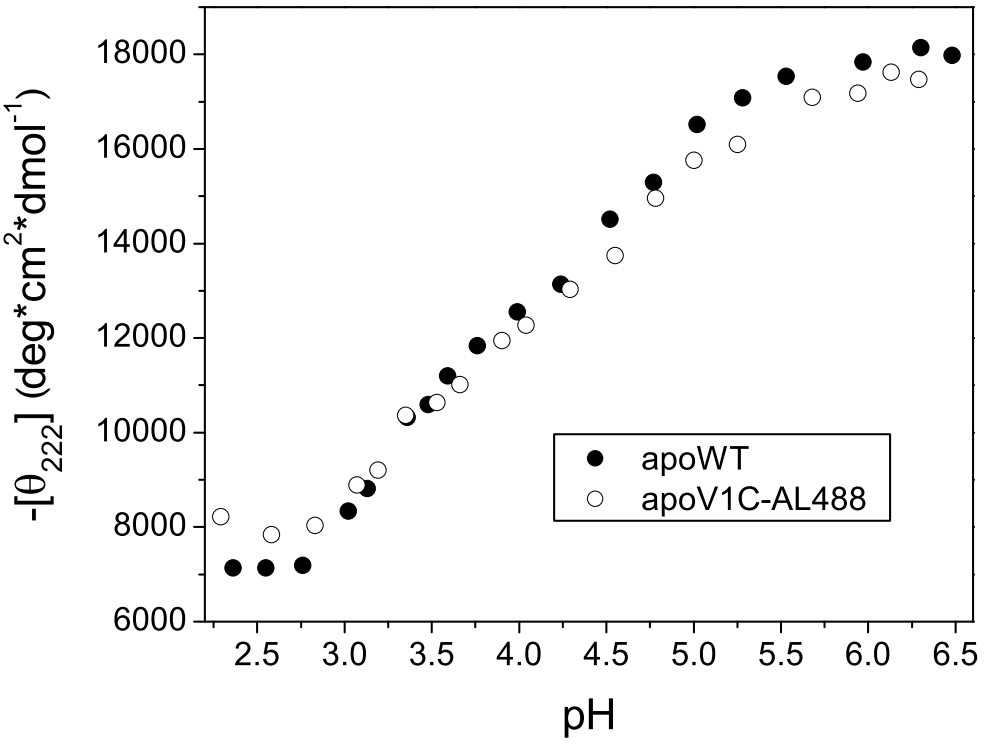

Figure 2.2: Acid denaturation of wild-type apoMb and Alexa 488-labeled V1C-apoMb (apoMb-V1C-AL488), monitored by circular dichroism at 222 nm at 20°C.

2.3.2 Measurement of conformational fluctuations

FCS measurements were made of apoMb V1C-AL488 in solutions ranging from pH 6.3 to 2.6. Figure 2.4 compares the correlation curves at three representative pH values with G(τ) normalized at G(τ = 200 s)=1 to emphasize the relative change in the shapes of the curves as a function of pH denaturation. The data shown in Figure 2.4 were chosen to represent the three stable conformations of

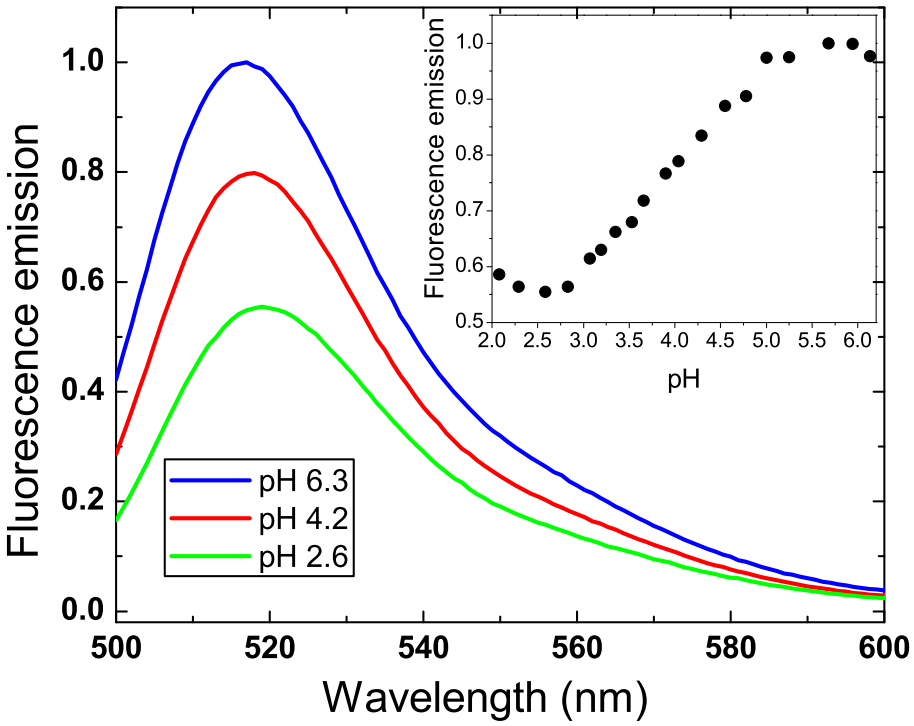

Figure 2.3: Bulk fluorescence emission of apoMb-V1C-AL488 in three representative pH solutions: pH 6.3 (red), pH 4.1 (blue), pH 2.6 (green). The curves are normalized using the absorption spectra of the solutions to correct for changes in Alexa 488 absorption at low pH and small differences in protein concentration. The inset shows acid denaturation of apoMb-V1C-AL488 as measured by Alexa 488 emission.

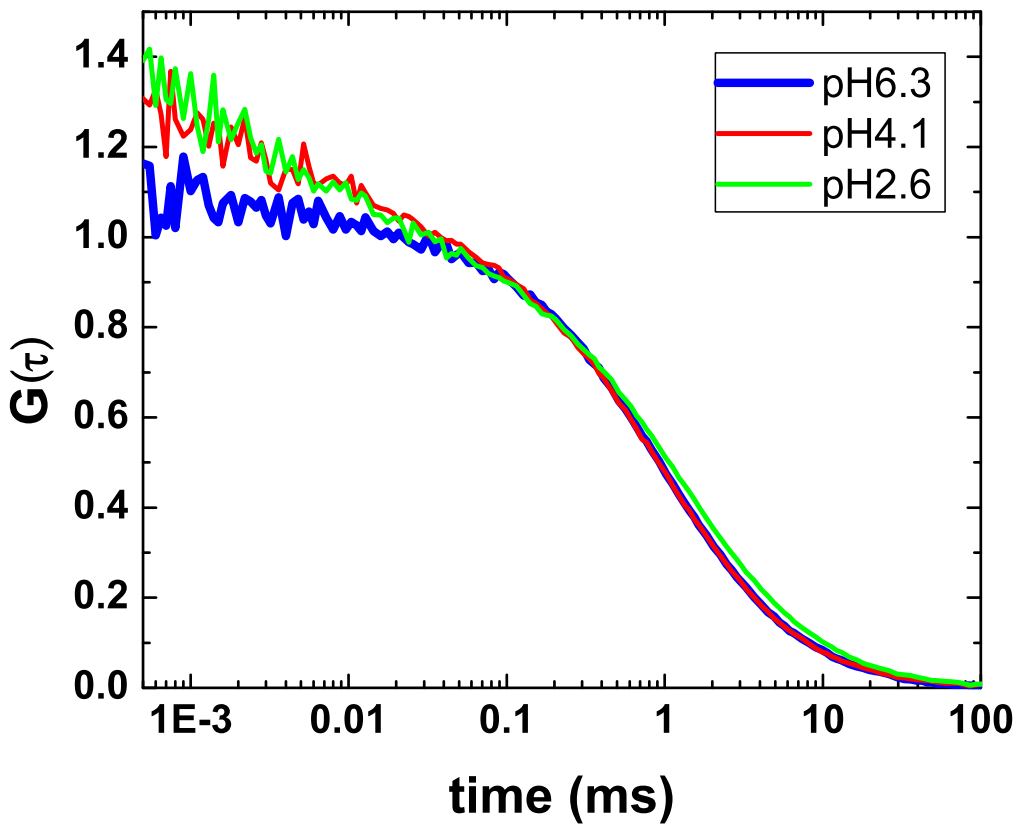

Figure 2.4: Correlation curves of apoMb-V1C-AL488 in three representative pH solutions: pH 6.3 (red), pH 4.1 (blue), pH 2.6 (green).

apoMb: N (pH 6.3), I (pH 4.1) and U (pH 2.6). The obvious differences amongst these curves indicate that the Alexa 488 probe is sensitive to differences in conformation of the protein. Above ≈0.3 ms, molecular diffusion dominates the correlation curves. The pH 6.3 and pH 4.1 correlation curves overlap at the longer times, indicating that I and N have comparable diffusion times. Unfolded apoMb (pH 2.6) exhibits a longer diffusion time, which suggests an ex-

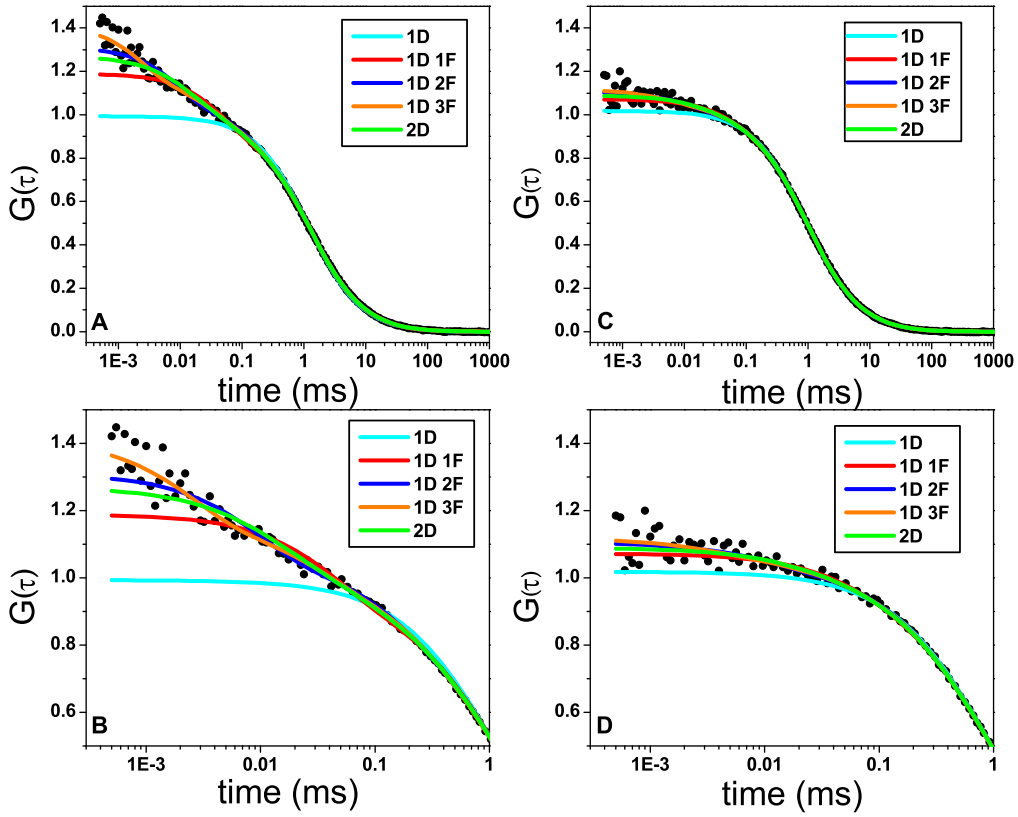

Figure 2.5: Correlation curves at pH 2.6 (panel A) and pH 6.3 (panel C), showing fits from the five different models considered: 1D (cyan), 1D1F (red), 1D2F (blue), 1D3F (orange), and 2D (green), where D designates a diffusion component and F designates a relaxation time component in the fits. Panel B and panel D are enlargements of panel A and panel C, respectively, to better illustrate deviations of each model at short timescales.

pansion of hydrodynamic radius. Below ≈0.3 ms, G(τ) largely reports on conformational fluctuations. U and I undergo similar fluctuations in this time regime, whereas some of these motions appear to be damped in N.

Data were fitted with five models seeking optimum representations: diffusion only with one or two diffusing components (1D and 2D) and with diffusion plus up to three kinetic components (1D+1F, 1D+2F, and 1D+3F) (see Section 2.2.4 for a complete discussion of models). The best fits from the five models for measurements at pH 2.6 and 6.3 are shown in Figure 2.5a and Figure 2.5c, with expanded axes in Figure 2.5b and Figure 2.5d, respectively.

It is clear from visual inspection that 1D and 1D+1F do not adequately fit the data at either pH. Only the 1D+2F and 1D+3F models return reasonable diffusion times and fit the data with flat residuals. To differentiate between the fits, the χ^2 values of the 1D+3F fit were subtracted from the χ^2 values of the 1D+2F fit for each pH (see Appendix A). At pH 2.6, 1D+2F and 2D fit the data poorly as well, leaving 1D+3F as the simplest model to accurately fit the data (Figure 2.5b). At pH 6.3 it is more difficult to discern goodness of fits between 1D+2F, 1D+3F, and 2D (by inspection as well as by residuals). 2D returns two diffusion times: $\approx$950 μs, which is $\approx$5% faster than what is expected for native apoMb in our system, and $\approx$30 μs for a second species. Though longer diffusion time could possibly correspond to apoMb and the presence of free dye could provide a second, faster diffusing component, the diffusion time of free Alexa 488 on our apparatus is $\approx$260 μs, thereby excluding 2D as a possibility at pH 6.3 as well. While it appears that 1D+3F may slightly improve the fit over 1D+2F at very short timescales (Figure 2.5d), this improvement does not justify the more complex model.

As for the entire pH series, between pH 6.3 and 4.35, the difference in χ^2 values between 1D+2F and 1D+3F is negligible. However, between pH 4.1 and 2.6, the χ^2 values jumped abruptly, indicating that the addition of a third exponent

considerably improved the fit. We therefore used the parameters extracted from 1D+2F for pH≥4.35 and those from 1D+3F for pH≤4.1 to interpret our results.

Although the amplitudes (A_i for pH≥4.35; B_i for pH≤4.1) of the correlation signals that can be attributed to structural fluctuations change as a function of pH (Figure 2.6), the relaxation times themselves are relatively constant (Figure 2.7). In the pH≤4.35 samples the amplitudes of the two relaxation times exhibit similar behavior; A_1, associated with the ≈100 μs fluctuation, grows from 3% to 10% as the pH decreases from 6.3 to 4.35, whereas A_2 reflects a ≈8 μs fluctuation that grows from ≈7% to ≈12%. In the pH≤4.1 samples, B_1 is associated with a ≈200 μs relaxation time and remains constant at ≈10%. The amplitudes of the two faster components in these samples are larger than B_1, and both increase in value as the pH decreases; B_2 (≈30 μs fluctuation) increases only a small amount, from ≈11% at pH 4.1 to ≈13% at pH 2.6, whereas B_3 (≈3 μs fluctuation) increases from ≈11% to ≈17% over the same pH range.

In choosing the number of exponential time decay components appropriate for each pH, we aimed to use the simplest model possible that accurately fit the data within the experimental uncertainties. We observed approximately an order of magnitude difference between each of the relaxation times we are able to assign. However, we cannot exclude the possibility that there are additional relaxation times that we do not observe, either because they are faster than the resolution of our measurements, or because they are not spectrally separable from the dominant time scales that we do observe.

To illustrate the pH dependence of the total contribution of all conformational dynamics to the correlation function, we sum the amplitudes of the non-diffusional components at each pH (Figure 2.8). The resulting curve is sigmoidal

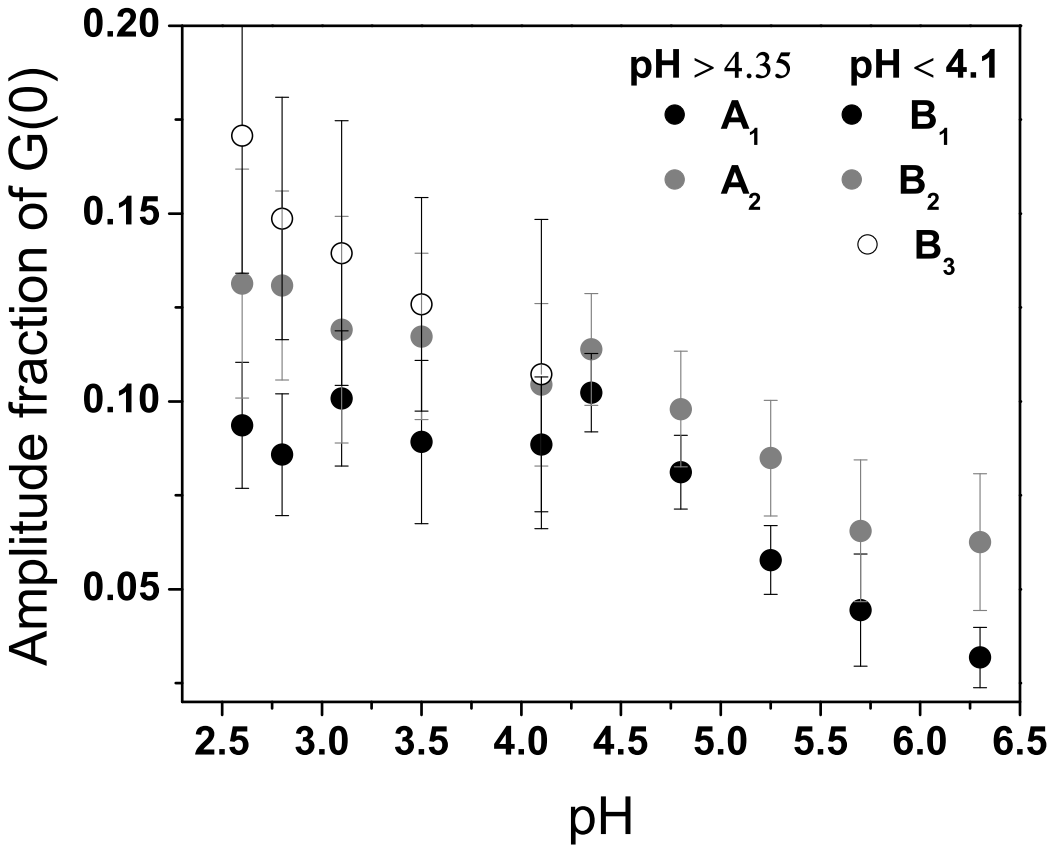

Figure 2.6: Best-fit values of structural fluctuation fractional amplitudes. pH dependence of the amplitudes extracted from 1D+2F (A_1 and A_2) and 1D+3F (B_1, B_2 and B_3) fits at pH≥4.35 and pH≤4.1 respectively; A_1 and B_1 (black circles), A_2 and B_2 (grey circles) and B_3 (unfilled circles). Each amplitude represents the prefactor of its associated exponential relaxation time to the overall correlation curve.

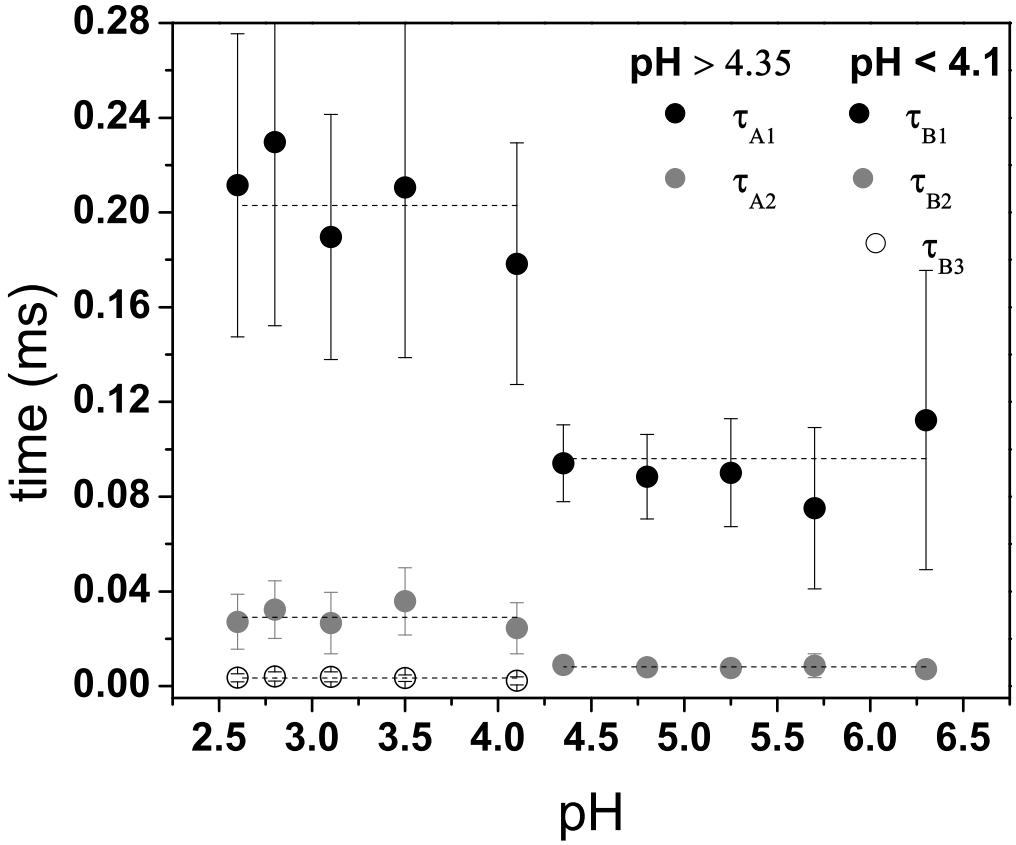

Figure 2.7: Best-fit values of structural fluctuation relaxation times. pH dependence of the relaxation times as a function of pH: τ_{A1} and τ_{B1} (black circles), τ_{A2} and τ_{B2} (grey circles), τ_{B3} (unfilled circles). Assuming that each τ does not change significantly across a relevant pH range, the dotted lines represent the average value of each τ across that range.

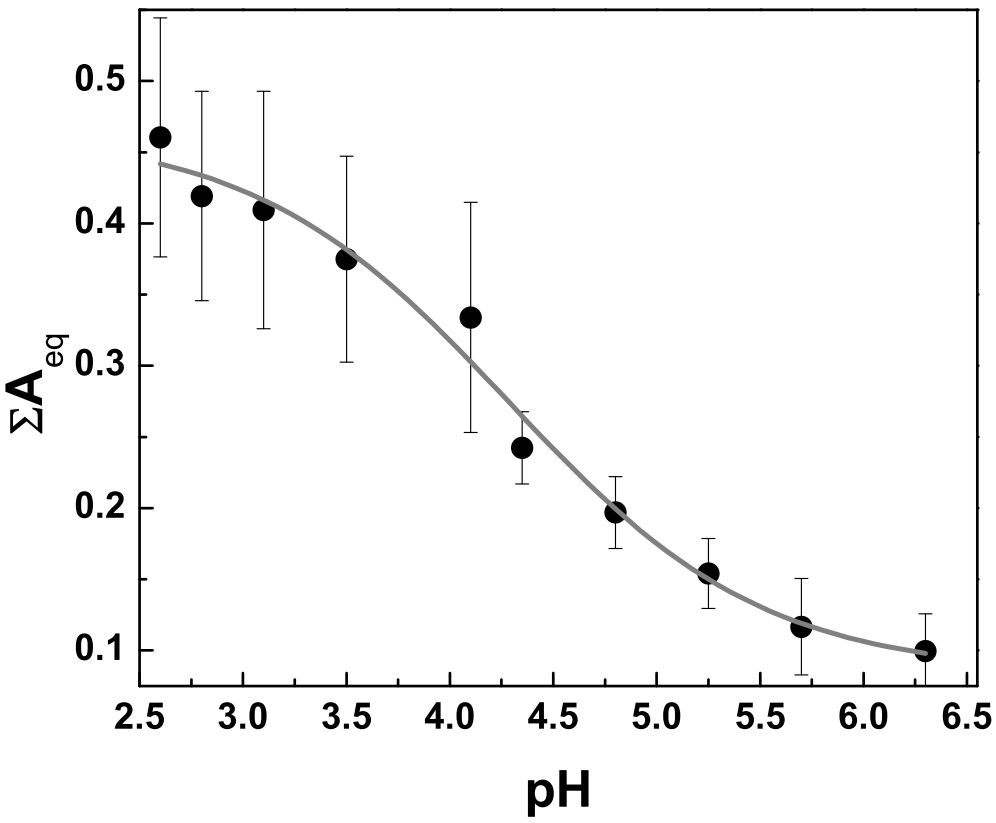

Figure 2.8: Normalized sum of the amplitudes of all conformational (non-diffusional) fluctuations, A_i and B_i, at each pH.

and exhibits a broad transition centered about pH 4.2. This trend reflects the substantially increasing extent of fluctuations as apoMb unfolds from N to I to U.

2.3.3 Measurement of protein diffusion as a function of folding

These measurements also provide diffusion times (τ_D) of the protein, which are directly proportional to the hydrodynamic radii (r_H) and represent an average over all fluctuating structures at each pH. τ_D can be extracted accurately from the correlation function fits; these values are plotted as a function of pH in Figure 2.9. The fitted diffusion time for the folded state at pH 6.3 is 1±0.015 ms, which corresponds to a hydrodynamic radius r_H≅2.0±0.1 nm. The r_H at pH 2.6 can be scaled from Figure 2.9 and compared with previous literature reports on published apoMb hydrodynamic radii [106]. ApoMb V1C-AL488 maintains a native-like conformation between pH 6.3 and 4.35, but shows a continuous increase in rH beginning at pH≈4.1. This increase is consistent with the I to U unfolding transition observed by CD (Figure 2.2). τ_D is relatively insensitive to the model used to fit the correlation function, and similar values for τ_D were found by using 1D+1F, 1D+2F, and 1D+3F. Thus, τ_D begins to increase between pH 4.35 and 4.1, regardless of the fitting model. This observation suggests that the conformational fluctuations of the molecule that result in the kinetics we observe are also changing between pH 4.35 and 4.1, providing a physical basis for our finding that two exponential components are sufficient to fit the data above pH 4.35, whereas three are necessary at lower pHs. Furthermore, we do not expect to find the same dynamics in pH≤4.1 solutions, where the protein explores the U and I conformations, as we do in the pH≥4.35 solutions, where the protein samples the more structurally similar I and N.

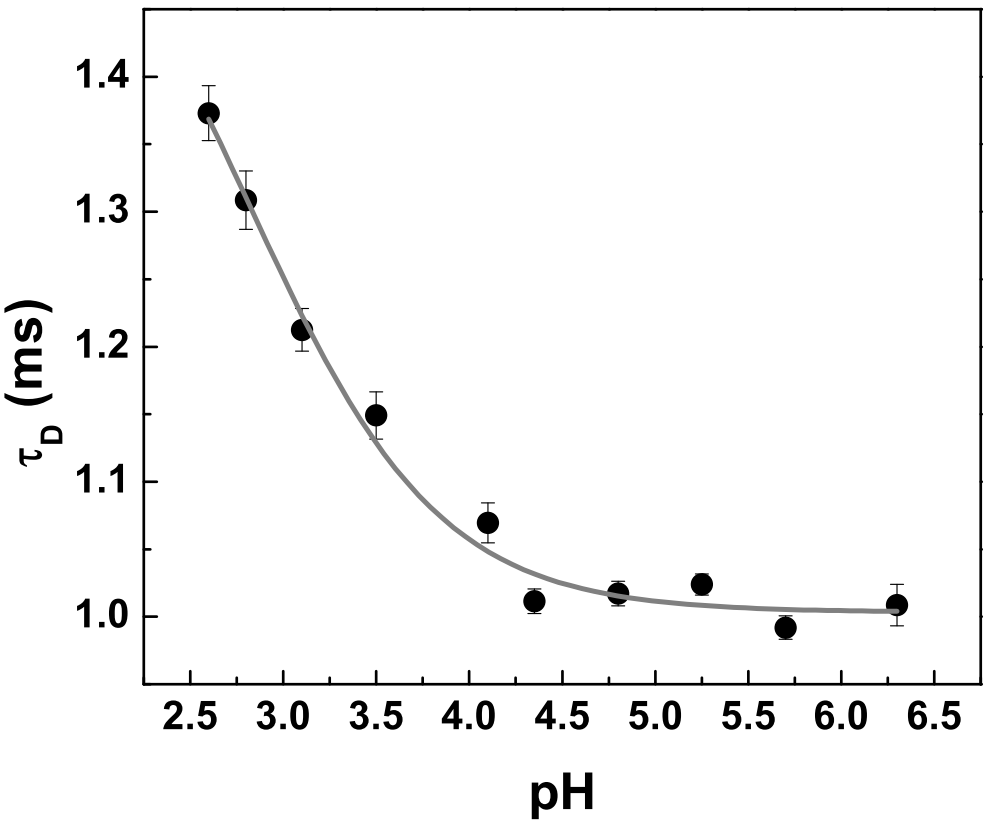

Figure 2.9: Plot of the diffusion time, τ_D, extracted from the correlation curve as a function of pH.

2.4 Discussion

Our results present clear evidence for structural fluctuations of apoMb across several distinguishable characteristic time scales: <10 μs (at all pH values), ≈30 μs (pH≤4.1), ≈100 μs (pH≥4.35), and ≈200 μs (pH≤4.1). We will first consider how the amplitudes and relaxation times from the fit functions relate to dynam-

ics of the protein chain, and then explore the origins of the quenching mechanism enabling observation of the protein dynamics.

2.4.1 Conformational dynamics of ApoMb folding

What physical processes are responsible for the fluorescence fluctuations observed by FCS? There are two general scenarios. First, the fluctuations may correspond to a two-state interconversion between folded and U conformations. In this case, the observed relaxation time would be predicted to exhibit a chevron-like pH dependence, with a maximum near the apparent pK_a of the transition, because the folding rate (which dominates at pH>pK_a) typically increases as pH increases, and the unfolding rate (which dominates at pH<pK_a) usually increases as pH decreases. Alternatively, the fluorescence changes may not arise only from the folding reaction itself, but also from localized dynamic quenching processes because of fluctuations within the partly folded or U states. Here, the relaxation time might be expected to remain relatively constant with pH, unless the motions that give rise to quenching are themselves constrained by pH-dependent structure.

Because a three-component model was favored to fit the FCS data below pH 4.1 and a two-component model above pH 4.35, we consider the two pH regimes separately. ApoMb-V1C-AL488 undergoes the I to N folding transition between pH 4.35 and 6.0. The N state is formed with a relaxation time (300 to 400 ms at pH 6) [41] that is much slower than the diffusion time of the protein, ≈1 ms (Figure 2.8). Therefore, FCS measurements in this pH range can be assigned primarily to fluctuations within I and/or within N, and less to their

interconversion. Superposition of the spectra of two relaxation times (τ_{A1}=100 μs and τ_{A2}=8 μs) adequately describes the fluctuations in this pH range. The amplitudes of both components decrease as the pH increases, suggesting that they primarily reflect conformational dynamics present in the I. Nevertheless, they persist in the N state, suggesting that the "folded" protein is sampling conformations associated with the intermediate state.

Our observations of conformational fluctuations in native apoMb are consistent with numerous reports in the literature that have characterized equilibrium dynamics. NMR studies have shown that at pH 6, a ≈20-amino acid stretch from the EF loop to the beginning of the G-helix (residues 82101) fluctuates between the folded state and unstructured or partially structured states [22]. An earlier investigation [60] that examined the effect of mutations on the native and acid denatured states of apoMb concluded that native apoMb fluctuates between several conformations. A third study [94] confirmed that finding by using energy transfer between the C-terminal Tyr and the A-helix Trps in horse apoMb (which has similar stability as sperm whale apoMb studied here) to show that a surprisingly large fraction of N state molecules sampled extended conformations.

The U to I transition occurs between pH 2.6 and 4.1. The fluctuations observed in apoMb-V1C-AL488 in this pH range can be described by a superposition of the broad spectra of three relaxation times: τ_{B1}=200 μs, τ_{B2}=30 μs, and τ_{B3}=3 μs. Jamin and coworkers [41] estimated the relaxation time of U/I folding to be 250 μs at 4°C and pH 4.2. The same study reported that the rate increases by a factor of 10 at 20°C, the temperature used in our experiments. These times are consistent with our measurements. However, the relaxation

times monitored by FCS are relatively constant from pH 2.6 to 4.2. They do not exhibit a maximum near pK_a (pH 3.4) that would be expected for a pH-dependent conversion between folded and U states. This finding suggests that the observed fluorescence fluctuations are caused by dynamic quenching by fluctuations within conformational states U and I, rather than to interconversion between them. The amplitudes of the two faster fluctuations (B_2 and B_3) decrease as pH is raised from 2.6 to 4.1, indicating that they reflect motions primarily within U. In contrast, the amplitude of the slowest phase (B_1) remains constant over this pH range, signifying that it represents a motion present in both U and I.

It is not surprising that the U state of the protein shows dynamic fluctuations of larger amplitude than the intermediate or N states. Indeed, broadening in NMR spectra of the acid denatured state of apoMb specifically predicts μs to ms contacts between the N-terminal portion of the protein, where Alexa 488 is located, and a stretch of amino acids that makes up the G-helix in the folded protein [23]. More specifically, a second NMR study found native-like interactions between the N-terminal A-helix region and the C-terminal GH-helix region of acid unfolded apoMb, indicating the transient formation of compact states in the denatured protein [59].

Contact rates for unstructured polypeptides have been successfully measured by using triplet-triplet energy transfer between exogenous probes [4] or naturally occurring amino acids [54, 55]. Two different studies observed that in longer polypeptide chains (>20-30 amino acids) the contact rates approached those predicted by polymer theory [27, 104]. We can thus consider the fluctuations seen in the U state of the protein in the context of their findings. Using

the data from these two studies [54, 51], we extrapolated the rates measured for model polypeptides to the most extreme case for apoMb, where the probe and quencher are separated by the entire length of the protein chain, to calculate contact times ranging from 400 ns to 6 μs, depending on the sequence of the model polypeptide. The upper end of these estimates is of the same time scale as the fastest of the fluctuations, $\tau_{B3} \approx 3$ μs, that we observe in the U state, and may represent the first contact time between distant parts of the protein chain [4]. However, we also observe relaxation times nearly two orders of magnitude slower than τ_{B3}. Thus, it is unlikely that the freely jointed Gaussian chain model, even when it takes into account excluded volumes and intrachain interactions, can account for the relatively slow ($\leq$200 μs) relaxation times that we measure. Rather, our findings of slow dynamics emphasize the importance of concerted motions of the protein chain, even in the U states, where secondary and tertiary contacts are thought to be transitory.

Other recent reports in the literature also report a range of time scales for chain dynamics of denatured proteins. In one, single molecule fluorescence was used to measure a chain reconfiguration time of $\approx$20 μs for ribonuclease H1, a value in good agreement with at least one of the time scales we observe [52]. However, a second single molecule study reports a reconfiguration time of $\approx$50 ns for a small cold shock protein (CSP), that becomes slower as the protein folds [75]. Although observations of relaxation times this rapid are not within the time resolution of our measurements, CSP is also considerable smaller (66 amino acids) than apoMb and is known to fold by a two-state mechanism, which may account for its simpler, more rapid chain dynamics.

2.4.2 Quenching Mechanism

The physico-chemical origins of the fluctuations in the fluorescence signal of apoMb-V1C-AL488 are attributed to interactions of the N-terminal Alexa 488 label with amino acids at various parts of the protein chain that quench the fluorescence emission of the dye. Because quenching is a short-range interaction (>2 Å) [53], we assume that quenching events occur via van der Waals contact between the dye and quencher on a time scale faster than is resolved in these measurements [77]. Although the specific quenching mechanisms are not the focus of this study, we did determine which amino acids are most likely to be involved. Trp, Tyr, His, and Met were identified as potential collisional quenchers of Alexa 488 by fluorescence quenching studies of the dye in solutions containing individual amino acids (see Chapter 3). Trp and Tyr are the most efficient quenchers. ApoMb contains two Trp residues (in the A-helix), and three Tyr residues (in the G-helix and near the C terminus) (Figure 2.1). The proximity of the Alexa 488 probe to the two Trp residues in the N state raises the possibility that it is the "flopping" of the dye (which has a long linker) to contact the A-helix that leads to some of the dynamics we observe. However, fluctuations of this type are expected to be on the order of nanoseconds or even shorter [4], significantly faster than detectable in our measurements. Although Met and His are less efficient quenchers than Trp and Tyr, the large number of His residues (12 His in the first 120 amino acids of apoMb as shown in Figure 2.1) implies additional weak quenching. Because multiple fluctuating components are observed on a broad range of time scales, it is most probable that multiple Alexa 488-residue interactions are responsible for the complexity.

Although our data do not specifically identify which contacts give rise to

the observed fluctuations, we may reason from NMR data to consider some of them. The fastest fluctuations (pH≤4.35, τ_{A2}≈8 μs; pH≥4.1, τ_{B3}≈3 μs) may represent contact between Alexa 488 and a proximal amino acid. In the acid-unfolded state (pH 2.6), the A-helix is known to form compact structures [22]. The A-helix contains the two Trps in close proximity to the fluorophore, and it is likely that contact made with these quenchers is reflecting the dynamics of the A-helix as it fluctuates between a random and a partially ordered structure. The decreasing amplitudes associated with τ_{A3} and τ_{B3} as the pH increases may be caused by increased stability of the A-helix in the I and N states or decreased access of the Alexa 488 to the Trp quenchers in these more compact structures. The remaining fluctuations are significantly slower (pH≤4.35, τ_{A1}≈100 μs; pH≤4.1, τ_{B1}≈200 μs and τ_{B2}≈30 μs), perhaps slow enough that they may reflect concerted chain motions. In the I state, the A-, G-, and H-helices are present and packed in a native-like core, with restricted flexibility as compared with the remainder of the protein [23]. The slow fluctuation found in solutions pH≤4.35 (τ_{A1}) is probably motion of the chains outside of this stable core region. As the protein folds, and the remainder of the protein forms a more stable conformation, the amplitude (A1) of this fluctuation decreases to only 3%. Likewise, the 200 μs (τ_{B1}) fluctuation is also likely to reflect a motion outside of the A-G-H core region (although not necessarily the same motion); its amplitude is constant across the pH≤4.1 range, indicating it is not affected by the formation of the core as the pH increases from 2.6 to 4.1. Both fluctuations (τ_{A1} and τ_{B1}) could be caused by contact of the N terminus with one or more of the His residues, the majority of which are located in the BF-helices.

A number of studies using FCS characterize protein fluctuations; however, each of those studies has shown relatively simple, single exponential kinetics

[9, 8, 93, 113]. In the case of cytochrome c, the quenching mechanism is between the fluorophore and the heme group of the protein, and thus single exponential kinetics might be expected [93, 113]. In a study by Chattopadhyay et al. [8], a protein was labeled with two fluorophores and monitored for self-quenching of these fluorophores, again anticipating a single mechanism that may be expected to give rise to a single relaxation time. In their earlier study, a single fluorophore was used, as in this study; nevertheless, only a single fluctuating component was observed. The more complex kinetics that we observe are undoubtedly caused by multiple interactions from complex chain motions that result in quenching of Alexa 488.

The correlation function we used to fit the data assumes a model with some number of independent quenching events. It is possible that choosing a different model will result in a different analytical form of the correlation function. Here, we have cautiously used the simplest models appropriate to interpret our data. Because the differences in time scales between each of the relaxation times we observe at each pH are roughly an order of magnitude, we have confidence in the broad kinetics we observe and resolve with our multivariate quenching mechanism. We have also enabled measurement of more complicated kinetics than have been previously accessible by our use of an expanded focal volume geometry for longer diffusion times (≈1 ms) of the protein, thus extending our sensitivity to longer time scales of structural fluctuations. In a smaller focal volume, the structural fluctuations τ_{A1} and τ_{B1} (≈100-200 μs) approach the protein diffusion time (≈150-250 μs) and thus are not separable from diffusion. The faster fluctuations we measure (τ_{A2}, τ_{B2}, and τ_{B3}) are similar in time scale to those measured in the above studies. It is appropriate to repeat the caveat that our choice of fitting the observed fluctuation correlation data with the smallest

number of exponential decay times sufficient to fit the data represents a distribution of fluctuations consisting of a superposition of fluorescence quenching events.

2.5 Conclusion

Remarkably complex, dynamic information about the broadband conformational fluctuations of the archetypal protein apoMb has been extracted without perturbing the system equilibria. As expected from other studies, the measured diffusion times τ_D and corresponding hydrodynamic radii of apoMb are found to expand by $\approx 40\%$ in the acid denatured state from the folded state. Of greater importance, we have found and have been able to characterize structural fluctuations of apoMb over approximately two orders of magnitude in time scale, as they vary with the folded, molten globule, and denatured states. Four distinguishable exponential relaxation times represent the broad conformational fluctuation spectra, revealing a general trend in slowing of the fluctuations concurrent with an increase in their associated amplitudes upon denaturation of the protein. This research has demonstrated that FCS can be a powerful tool for the study of protein folding associated conformational fluctuations. Future directions aim to distinguish the complex behavior observed here. These studies may benefit from use of double-labeled protein with a specific quenching mechanism, as well as mutations of amino acids thought to be responsible for quenching. It may be particularly interesting to monitor by FCS the fluctuations associated with the known transitions of apoMb from an assembly of α-helices to reported β-ribbon structures and aggregates that have been observed under aggressive conditions [25, 12].

CHAPTER 3

MECHANISMS OF QUENCHING OF ALEXA FLUOROPHORES BY NATURAL AMINO ACIDS*

3.1 Introduction

Quenching of fluorophores by the same proteins to which they are linked is a phenomenon that is neither well-known nor well-characterized. Fluorescence spectroscopy has recently emerged as an important tool in studying protein conformational fluctuations and protein-protein interactions. Extrinsic fluorescent molecules, with reactive side groups, are covalently attached to specific labeling sites within the proteins of interest, and changes in fluorescence intensity or lifetime of these fluorophores can reflect on changes in their local environment. Mechanisms like Förster Resonance Energy Transfer (FRET), which will be explained in more detail in Chapter 4, can allow for precise measurement of the distance between acceptor and donor fluorophores attached to various parts of a protein and indirectly report on the conformational changes within the protein [102]. However, the effect of the target protein itself on the attached fluorophore has to be considered when making accurate quantitative measurements of fluorescence intensity. In this chapter, we elucidate the quenching effects that the protein itself can have on the fluorophore attached to it, which can also help in selecting the optimal labeling site in a protein.

After labeling, proteins are usually assayed to ensure that the attached fluo-

*Originally published as: Chen, H., Ahsan, S. S., Santiago-Berrios, M. B., Abruna, H. D., Webb, W. W. (2010). "Mechanisms of quenching of Alexa fluorophores by natural amino acids," Journal of the American Chemical Society 132(21), 7244-7245.

rophore(s) does not affect their intrinsic structure and function. However, it cannot be assumed that the amino acids in close proximity to the attachment site do not change the fluorophore's photophysical properties. Spacers, such as a short saturated carbon chain, can be added between the fluorophore and the attachment site to mitigate unwanted interactions. However, these spacer molecules are often flexible and the fluorophore could still come into contact with amino acids near the attachment site. Indeed, it has been observed that attached fluorophores can be quenched by contact with amino acids within the same molecule, and this property has been exploited to report on changing conformational states [10, 72] or intramolecular dynamics of proteins [110, 66, 17]. For example, this effect has been used to measure the chain dynamics in Alexa488-labeled apomyoglobin using Fluorescence Correlation Spectroscopy (FCS) as discussed in Chapter 2 [10]. While the specific mechanism of this quenching phenomenon is unknown, it has been attributed to photoinduced electron transfer (PET). Broadly, photoinduced electron transfer describes the phenomenon where a molecule (donor) excited by light transfers its excited state electron to another molecule (acceptor) when they are in close van der Waals contact.

The Alexa dyes from Invitrogen^{TM}are a family of fluorophores commonly used to label proteins for both in vitro and in vivo experiments because of their superior photophysical properties. These rhodamine-derived fluorophores are more stable to photobleaching or conversions to the triplet states [81], making them popular choices for fluorescence imaging. Alexa dyes have high quantum efficiencies, thus making them suitable for single molecule studies where brightness is an important factor in the choice of fluorophore. Alexa pairs like Alexa 488/Alexa 555 and Alexa 488/Alexa 594 are often used as conformational reporters in FRET experiments.

In this chapter, we identify amino acids that quench the fluorescence of Alexa 488, Alexa 555 and Alexa 594. For Alexa 488, we identify as many as four quenchers - Tryptophan (Trp), Tyrosine (Tyr), Histidine (His) and Methionine (Met). We observe a combination of static and collisional mechanisms in the quenching of Alexa 488 by these amino acids.

3.2 Theory of contact quenching

Fluorescence quenching can be broadly classified as arising from either dynamic or static mechanisms [53].

As shown in Figure 3.1, dynamic quenching arises due to molecular collisions between the excited state fluorophore and its quencher that occur during the excited state lifetime of the fluorophore, and result in relaxation back to the ground state without the emission of fluorescence. In this case, we expect the fluorescence lifetime of the molecule to be sensitive to the presence of the quencher. In the absence of the quencher, the fluorescence lifetime of the fluorophore τ_0 depends on the rate of decay from the excited state γ:

$$\tau_0 = \frac{1}{\gamma} \tag{3.1}$$

In the presence of the quencher, relaxation to the ground state depends on the rate γ as well as the rate k_D that is due to collisions with the quencher of concentration $[Q]$. In this case, the fluorescence lifetime τ is given by:

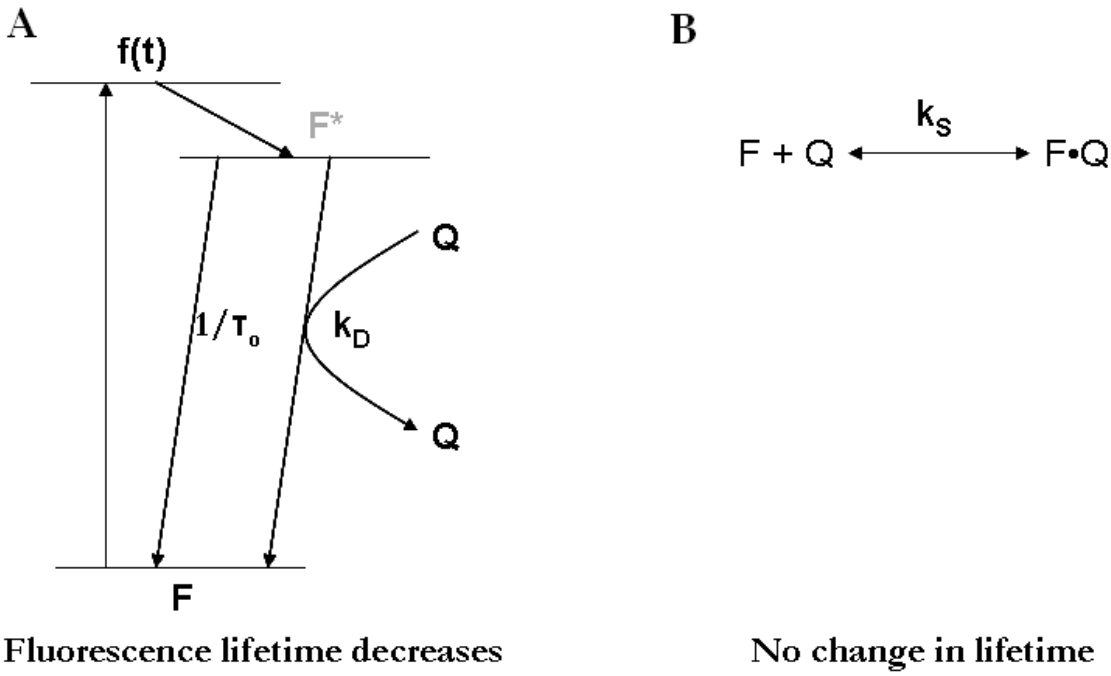

Figure 3.1: Two broad mechanisms can give rise collisional quenching when the fluorophore (F) and the quencher (Q) interact. (A) In dynamic quenching, collisions during the excited state lifetime of the fluorophore (F*) result in relaxation to the ground state without the emission of fluorescence. (B) In static quenching, formation of a non-fluorescent complex (F•Q) reduces the number of molecules that can excited for fluorescence.

$$\tau = \frac{1}{\gamma + k_D[Q]} \tag{3.2}$$

Then, the ratio of fluorescence lifetime in the absence and presence of quencher can be written as:

$$\frac{\tau_0}{\tau} = 1 + K_D[Q] \tag{3.3}$$

where K_D is the dynamic quenching constant and $[Q]$ is the concentration of the quencher.

Static quenching arises through the formation of a non-fluorescent complex between the two species, as depicted in Figure 3.1b. In this case, since quenching acts on the ground state, only the fluorescence intensity but not the fluorescent lifetime is expected to change in the presence of the quencher. The association constant between the fluorophore and the quencher can be given by:

$$K_S = \frac{[F \bullet Q]}{[F][Q]} \tag{3.4}$$

where $[F \bullet Q]$, $[F]$ and $[Q]$ are the concentrations of the nonfluorescent complex, the fluorophore and the quencher respectively. Since the total concentration of fluorophore $[F_0]$ is the sum of the fluorescent and non-fluorescent complexes, we can write the constant as:

$$K_S = \frac{[F_0] - [F]}{[F][Q]} \tag{3.5}$$

A simple rearrangement gives us an equation with a form similar to Equation 3.3.

$$\frac{[F_0]}{[F]} = 1 + K_S[Q] \tag{3.6}$$

$[F_0]$ is the total concentration of fluorophore and can be represented as the fluorescence intensity in the absence of quencher. $[F]$ is the concentration of the fluorescent species and can be represented as the fluorescence intensity in the presence of quencher. Then, we can write:

$$\frac{F_0}{F} = 1 + K_S[Q] \tag{3.7}$$

By measuring the fluorescence intensities and lifetimes in the presence and absence of quenchers, we can distinguish between the static and dynamic mechanisms.

3.3 Materials and Methods

3.3.1 Steady state fluorescence measurements

Fluorescence emission scans of Alexa 488, Alexa 555 and Alexa 594 hydrazides (Invitrogen Corp., Carlsbad, CA) were performed in the presence and absence of the 20 naturally-occurring L-amino acids. All experiments were performed on commercially bought L-amino acids (Sigma-Aldrich Corp., St. Louis, MO) and dissolved in 100 mM sodium phosphate buffer at pH 7. Steady state fluorescence experiments were performed using a spectrofluorimeter (PTI Inc., Birmingham, NJ). The fluorescence intensity at the peak emission wavelength was used to calculate to the amount of quenching in the presence of the amino acids. A plot of F_0/F as a function of the concentration of quencher was created and the slope was used to calculate the Stern Volmer quenching constants K_D and K_S.

3.3.2 Time-resolved fluorescence lifetime measurements

Fluorescence lifetimes of Alexa 488, τ, were measured using a homebuilt lifetime setup. Briefly, a Ti-Sapphire laser (Spectra-Physics, Newport Corp., Irvine, CA) was used for excitation at 895nm. The 80 MHz pulses were picked down to 10 MHz by a Pockels cell pulse picker (ConOptics Inc., Danbury, CT). The emission was separated from the excitation by a 700 nm short-pass filter and a 530/15 nm bandpass filter, and then detected by a microchannel plate photomultiplier tube (Hamamatsu Corp., Bridgewater, NJ). The signal was processed by a time-correlated single photon counter card (TCSPC) (Becker & Hickl GmbH, Berlin, Germany), and analyzed in the accompanying SPC Image software. The emission fluorescence was collected at the magic angle of 54.7° relative to the excitation to exclude the effects of polarization. The temporal response function of the instrumentation was measured using second harmonic generation from a potassium dihydrogen phosphate (KDP) crystal and was applied to correct all data sets. The resulting fluorescence decay curves were fitted with multi-exponentials and the precision of fits were determined by the reduced χ^2 values.

3.3.3 Electrochemical measurements of oxidation potentials

Cyclic voltammetry was performed in a phosphate buffer solution using glassy carbon electrodes against an $Ag^+/AgCl$ reference electrode and a platinum wire as a counter electrode, at a scan rate of 100 mVs_{-1}. Prior to measurement, the solution was purged with nitrogen gas for 15 min. All potentiometric measurements were made with a BAS-27 W potentiostat (Bioanalytical Systems Inc., West Lafayette, IN) using a WinDaq card and software as an acquisition method.

There is some uncertainty in these values since the redox processes of the amino acids are chemically irreversible processes. However, a reasonable estimate of this effect would be no more than about 100 mV.

From the measurements of oxidation potentials, we can determine the free energy of the electron transfer reaction, using the Rehm-Weller relation [91] given by:

$$\Delta G^o = E^o_{ox}(D/D^+) - E^o_{red}(A/A^-) - E_{0,0} + C \tag{3.8}$$

where $E^o_{ox}(D/D+)$ and $E^o_{red}(A/A-)$ are the first oxidation and reduction potentials of the electron donor (D) and acceptor (A) respectively, $E_{0,0}$ is the zero-zero transition energy of the fluorophore, and C is a Coulombic term that can be neglected in water.

3.4 Results

3.4.1 Quenching occurs through both dynamic and static mechanisms

To identify fluorescence quenchers of Alexa 488, Alexa 555 and Alexa 594, we performed fluorescence intensity measurements of the fluorophores in the absence (F_0) and the presence (F) of each of the 20 natural amino acids. We identified Trp as a common fluorescence quencher for all three fluorophores. Additionally, we observed that Tyr, His and Met also quench Alexa 488 fluorescence.

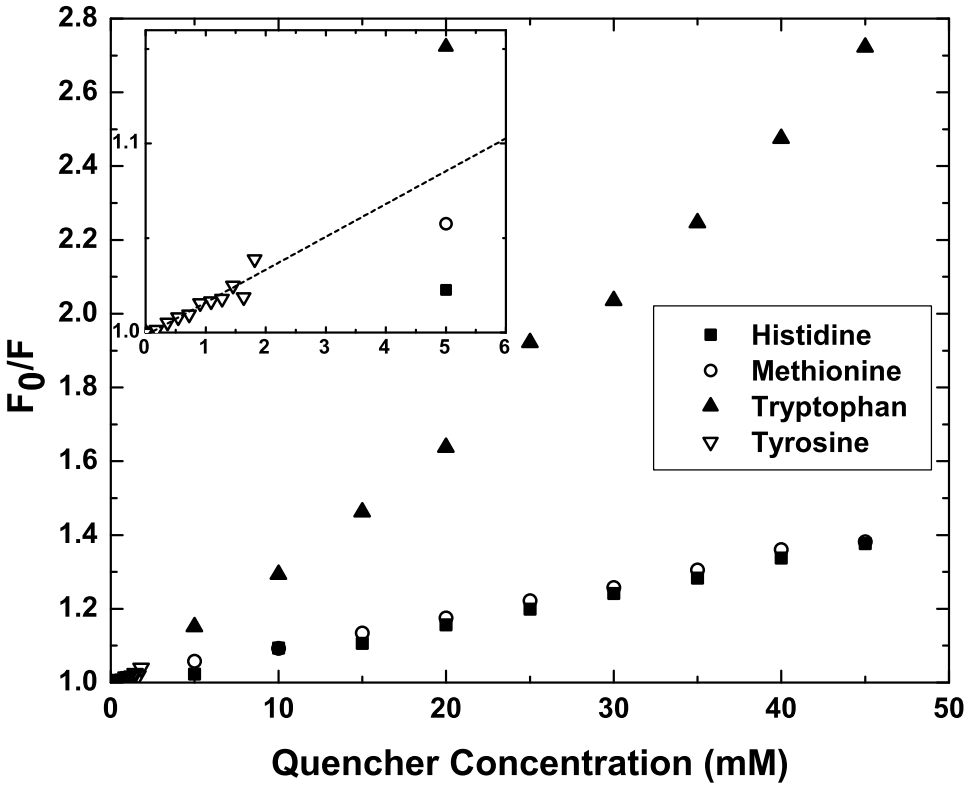

Figure 3.2: Comparison of quenching of Alexa 488 as measured by fluorescence intensity across four amino acids of interest. Tryptophan exhibits the strongest quenching effect, while Histidine and Methionine show similar quenching rates. Inset shows enlargement of low quencher concentrations up to 2 mM as Tyrosine is limited by its low solubility.

The fluorescence quenching due to each amino acid is depicted in the Stern-Volmer plot as shown in Figure 3.2.

To distinguish between the static and dynamic mechanisms of collisional quenching, we also performed fluorescence lifetime measurements of Alexa 488

in the absence (τ_0) and presence (τ) of each of the four amino acids identified as quenchers from intensity measurements. All fluorescent decays were well fit by a single exponential, from which we extracted the fluorescence lifetime. We measured a 4.0 ns fluorescence lifetime for Alexa 488 hydrazide in the absence of quencher, which corresponds well to previously published values of 3.9 ns [102] and 4.1 ns (data provided by Invitrogen Inc.). We plotted the lifetimes in the form of a Stern-Volmer plot for each amino acid and overlaid these with the same plots from intensity measurements for comparison, as shown in Figure 3.3.

To measure the feasibility of electron transfer between the species of interest, we performed cyclic voltammetry experiments. A typical scan is shown in Figure 3.4, where the first oxidation and reduction peaks of Alexa 488 are highlighted by the arrows. A summary of the first oxidation and reduction peak of Alexa 488 and the four amino acids is displayed in Table 3.1. The free energy of electron transfer between Alexa 488 and the quencher amino acid was calculated by the Rehm-Weller equation (Equation 3.8). While a favorable free energy shows that the reaction is thermodynamically feasible, we are not able to distinguish if the reaction will proceed kinetically. In the table, we show the more likely direction of electron transfer based on literature on the amino acids.

3.5 Discussion

Our results demonstrate clearly that Trp, Tyr, His and Met are quenchers of Alexa 488 fluorescence. While Trp and Tyr are known to quench certain fluorophores [110, 29] the inclusion of His and Met was surprising and led us to investigate the quenching of Alexa 488 in greater detail. Comparing the Stern-

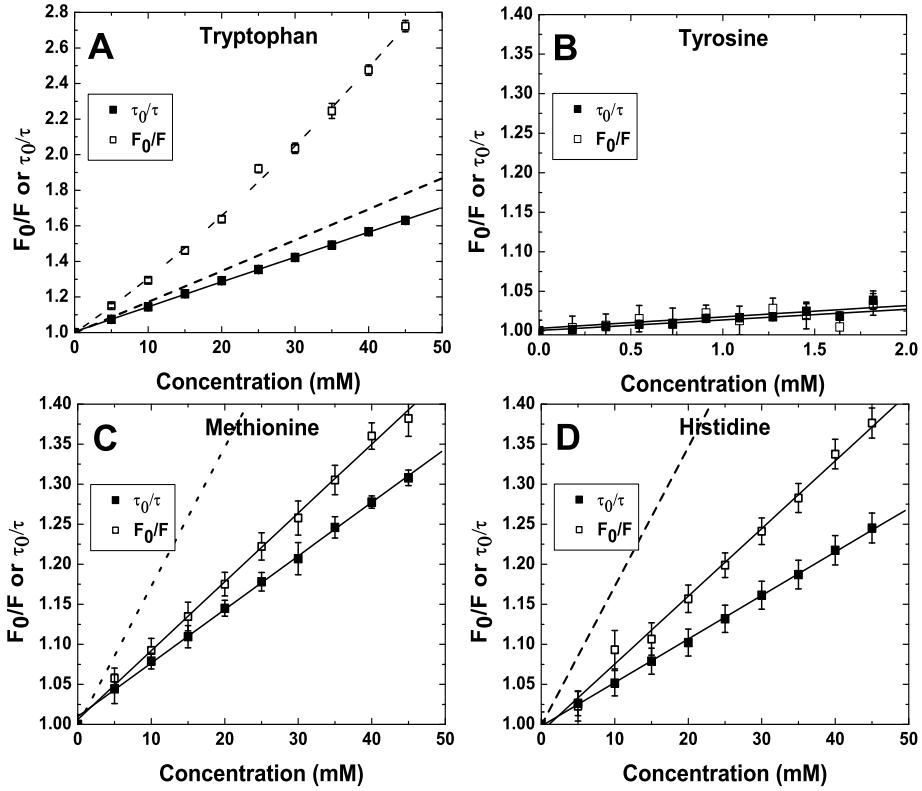

Figure 3.3: Stern-Volmer plots of Alexa 488 quenched by (A) Tryptophan, (B) Tyrosine, (C) Methionine and (D) Histidine. Open squares represent total quenching detected by bulk fluorescence intensity measurements (static and dynamic quenching) while shaded squares represent quenching detected by fluorescence lifetime measurements (dynamic quenching only). Solid lines represent linear fits to the data except for the bulk fluorescence data of Tryptophan where the static quenching component is significant and causes a deviation from linearity. Note that data for Tyrosine is limited to 2 mM of quencher concentration due to its low solubility. Dotted lines in A, C and D show the Tyrosine data from B for comparison.

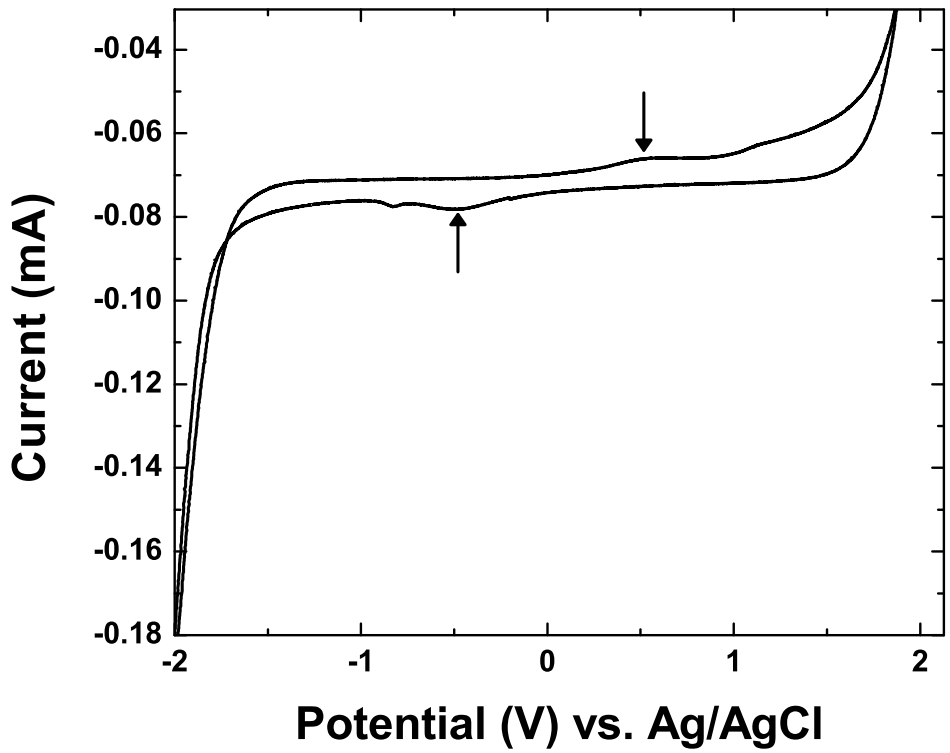

Figure 3.4: A typical voltammetry scan, in this case Alexa 488, showing the first oxidation and reduction potentials (highlighted by arrows).

Volmer plots of F_0/F vs. [Q] across all four amino acids in Figure 3.2, we observe that Trp and Tyr are strong quenchers of Alexa 488 fluorescence, while His and Met quench to a smaller extent. As shown in Figure 3.3, differences between the slopes of the lifetime and the intensity measurements indicate that quenching of Alexa488 by Trp, His and Met have both static and dynamic components, while that of Tyr occurs only through dynamic mechanisms. We will now discuss the mechanism of quenching due to each amino acid in detail.

	Oxidation Potential (V)	Reduction Potential (V)	Free energy (eV)	From/To Alexa 488
Alexa 488	-0.48	0.53	–	–
Trp	-0.45	0.79	-1.15	To
Tyr	-0.50	0.74	-1.38	To
Met	-0.46	0.72	-1.21	To
His	-0.45	0.53	≈ -1.4	To/From

Table 3.1: Reduction and oxidation potentials of Alexa 488 and amino acid quenchers vs. Ag/AgCl is used to calculate the free energy of electron transfer by the Rehm-Weller equation. While the free energy could be favorable in either direction of election transfer, the more likely direction based on known behavior of the amino acids is shown

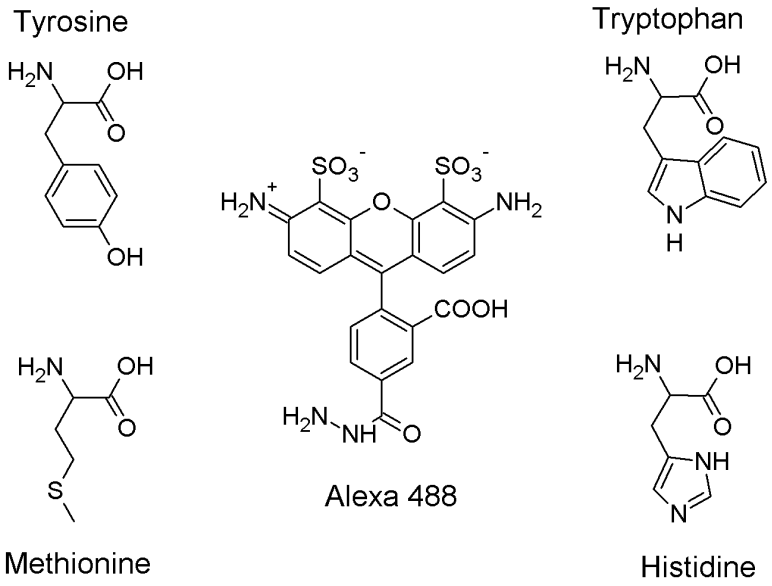

Figure 3.5: Chemical structure of Alexa 488 and four amino acids

3.5.1 Static quenching can be caused by ring stacking reactions

The difference between the slopes of the lifetime and the intensity measurements indicate that quenching of Alexa 488 by Trp has both static and dynamic components (Figure 3.3a), with a dynamic quenching rate of 3.5x10^9 $M^{-1}s^{-1}$. The predicted collision rate between Alexa 488 and Trp in solution is calculated from the Smulochowski equation as 9.3x10^9 $M^{-1}s^{-1}$. Hence, the fraction of total collisions that result in dynamic fluorescence quenching is 0.379±0.006. This value may indicate a dependence on the relative orientations (steric factor) of the rings of Alexa 488 and Trp when they collide (Figure 3.5). This agrees with observations that the quenching of the red-absorbing oxazine derivative dye MR121 by Trp occurs through photoinduced electron transfer (PET) mediated by ring-ring interactions [110]. Trp and Tyr have also been observed to undergo ultrafast electron transfer with riboflavin in flavoproteins [119].

We observed that the transfer of an electron from Trp to Alexa 488 is indeed highly favorable, with a free energy of -1.15 eV. From Figure 3.2a, we also calculated the static quenching constant K_S to be 15.1 M^{-1}. We note that approximately half of the total quenching by Trp arises from static mechanisms. Static quenching often arises when the fluorophore and quencher take part in stacking interactions and form a non-fluorescent complex [53]. This could possibly explain the ubiquitous quenching of different Alexa fluorophores by Trp.

3.5.2 Dynamic quenching is caused by Photoinduced Electron Transfer (PET)

Measurements of quenching by Tyr are limited by the low solubility of Tyr in aqueous solution. From fluorescence intensity measurements, we observe a strong quenching, comparable to Trp (Figure 3.2). The Stern-Volmer plot from lifetime and intensity measurements shows the same slope (Figure 3.3b), indicating that the quenching of Alexa 488 by Tyr occurs entirely through the dynamic mechanism. Since phenylalanine (Phe), which differs from Tyr only by the lack of the phenolic -OH group, does not quench Alexa 488 fluorescence, we surmise that quenching arises from PET between Alexa 488 and the -OH group on Tyr. Tyr is known to be a good electron acceptor that undergoes photoinduced electron transfer [29]. From cyclic voltammetry, we observe that there is indeed a favorable free energy of -1.38 eV for electron transfer from Tyr to Alexa 488.

3.5.3 Quenching mechanism of Methionine

We observe from the Stern-Volmer plot (Figure 3.3c) that Met quenches Alexa 488 primarily by collisional quenching, with a rate of 1.67x10^9 $M^{-1}s^{-1}$. The fraction of total collisions that result in quenching is 0.195±0.008. The side chain of Met contains a sulfur atom embedded in a saturated hydrocarbon chain, and since the sulfur atom is known to be susceptible to oxidation, we speculate that the quenching to arise during molecular collisions from PET originating from the reactive sulfur atom. This is feasible as electron transfer from Met to Alexa 488 has a free energy of -1.21 eV.

3.5.4 Quenching mechanism of Histidine

The quenching of Alexa 488 by His has both static and dynamic components (Figure 3.3d) with a dynamic quenching rate of 1.37x10^9 $M^{-1}s^{-1}$. The fraction of total collisions that result in quenching is 0.14±0.01. The imidazole group of His is known to undergo both oxidation and reduction reactions and this makes PET a possible mechanism of collisional quenching, as demonstrated by a favorable free energy of approximately -1.40 eV for electron transfer to or from Alexa 488. Similar to Trp, static quenching most likely occurs through stacking interactions between the imidazole ring of His and Alexa 488.

3.6 Conclusion

While it is often assumed that fluorophores do not interact with amino acids in the proteins they label, we show in this chapter that the fluorescence of Alexa 488, Alexa 555 and Alexa 594 is, in fact, quenched by interactions with Trp, Tyr, Met and His residues through a combination of static and dynamic quenching mechanisms. In light of this finding, more consideration should be taken in choosing labeling sites for fluorophores within proteins, especially when quantitative comparison of fluorescence brightness is important. Indeed, we have observed a significant dimming of Alexa 488 when the fluorophore is attached less than 3 residues away from a Tyr residue (data not shown). Thus, the potential effect of intramolecular quenching should be considered in the interpretation of data that involves quantitative measurements of fluorescence intensity.

CHAPTER 4

FLEXIBITILITES OF SINGLE-STRANDED NUCLEIC ACIDS MEASURED BY SINGLE MOLECULE FÖRSTER RESONANCE ENERGY TRANSFER (SMFRET)*

4.1 Introduction

RNA molecules in the cell are dynamic and undergo structural changes as they transmit and process genetic information. Because base-paired RNA is relatively rigid, dynamic processes require weakly bonded or disordered regions that confer flexibility to the overall structure. For RNAs like riboswitches that exchange between multiple structures in equilibrium [63], conformational disorder is often an intrinsic property of the molecule and important for biological function. Even relatively stable molecules like catalytic introns and transfer RNA must pass through a disordered phase while folding. Thus, progress toward a mechanistic understanding RNA folding and dynamics will require detailed knowledge of nucleic acid chain flexibility and its dependence on base content, solution conditions, and molecular context.

In light of its importance to biology, it is surprising that RNA flexibility has not been studied in as much detail as DNA flexibility. Despite the chemical similarity of the RNA and DNA backbone, there is ample evidence from x-ray crystallography that the identity of the sugar (ribose vs. deoxy-ribose) affects backbone conformations [74]. However, researchers have used the properties of DNA to understand RNA folding [96] because corresponding information for

*To be published as: Chen, H, Meisburger, S. P., Pabit, S. A., Sutton, J. L., Webb, W. W., Pollack, L. (2011). "The ionic strength-dependent persistence lengths of ssRNA and ssDNA,"

RNA was lacking. This difficulty motivates our present efforts to measure and directly compare the flexibilities of single-stranded nucleic acids (ssRNA and ssDNA).

In the cell, RNA and DNA interact with cations that screen the negatively charged phosphate backbone. Both diffuse and specifically bound ions are important for RNA folding [18], and divalent ions are almost always required to stabilize RNA tertiary structures [39, 14, 88, 7]. A full mechanistic description of these ion effects is complicated by the fact that ions can interact with RNA differently during various stages of folding [19, 116]. The interactions between ions and ordered nucleic acid structures have been characterized theoretically and experimentally by our group and others [79, 49, 1, 90]. However, a comparable analysis for the flexible and disordered regions is hampered by the lack of detailed knowledge of backbone conformations in solution, and how these depend on base content, sugar type, and salt ions present.

Previous measurements of flexible single-stranded nucleic acids using fluorescence correlation spectroscopy (FCS) [16], single molecule Förster Resonance Energy Transfer (smFRET) [56, 73] and force spectroscopy [15, 68, 97, 98] have not provided a consensus picture of backbone conformations. These studies use a variety of nucleic acid constructs (of varying length composition and length, with or without molecular handles or fluorescent tags) and interpret data with different models and starting assumptions. Equilibrium structural measurements of disordered macromolecules are inherently problematic when one tries to describe a large ensemble of structures with just a few experimental variables. This difficulty is found in the protein folding literature, where small angle X-ray scattering (SAXS) and smFRET measurements of denatured state

collapse appear irreconcilable [86, 40, 69]. Hence, there is a need for studies of disordered macromolecules that apply multiple complementary techniques to the same system and interpret these results using the same models and assumptions.

In this study, we provide a direct comparison of the properties of ssDNA and ssRNA in solution. A powerful combination of SAXS and smFRET reveals important structural information about dT_{40} and rU_{40}, nucleic acid chains that undergo minimal stacking interactions. The FRET efficiency and the full SAXS curve are consistent with the WLC model [50] for flexible polymers. We show that ssRNA and ssDNA chains differ in both flexibility and spatial extent. This difference is likely to influence their molecular recognition and folding properties and has the potential to further our understanding of the different roles played by DNA and RNA in cells. Using smFRET over a wide range of Mg^{2+} and Na^{+} concentrations, we find that the apparent charge screening efficiency of Mg^{2+} is anomalously large compared to Na^{+} in both ssRNA and ssDNA; comparable to previously measured ratios for double stranded DNA [90]. The strong interaction between Mg^{2+} and unfolded nucleic acids challenges our understanding of how ions influence the folding process.

4.2 Materials and Methods

4.2.1 Sample Preparation

All nucleic acid samples were synthesized and HPLC purified (Integrated DNA Technologies Inc. Coralville, IA). For smFRET experiments,

dT_{40} and rU_{40} were also labeled by the vendor with Cy3 and Cy5 fluorophores on the 3′ and 5′ ends respectively (Cy3-dT_{40}-Cy5 and Cy3-rU_{40}-Cy5). For experiments on the effect of an attached double helix, two separate strands (3′-ACCGCTGCCGTCGCTCCG-Cy5-5′ and 3′-Cy3-dT_{40}-CGGAGCGACGGCAGCGGT-5′) were mixed in 1:1 ratio in pH 8 20 mM Tris buffer and annealed for 2 min at 95°C ($dT_{40} - duplex$).

For SAXS experiments, dT_{40} and rU_{40} were resuspended in TE buffer at pH 7.5. Samples were buffer exchanged against 1 mM Na-MOPS at pH 7.0 with 100mM NaCl. The concentration of nucleic acids in each sample was determined using absorption at 260 nm (Cary 50, Varian Inc., Australia), and a concentration series (30, 75, and 150 μM) was made by dilution.

SmFRET experiments were performed in eight-chambered coverglass slides (Labtek, Nunc, Rochester, NY) that had been preincubated with 0.1% bovine serum albumin (BSA) solution for at least 2 hours and then washed with deionized water to remove the unbound BSA. smFRET experiments were performed in 20 mM Tris buffer at pH 8 with varying amounts of NaCl and $MgCl_2$ while the concentrations of the nucleic acids used were approximately 100 pM to ensure that at most only one molecule is in the focal volume at any time.

4.2.2 SAXS measurements

SAXS data were acquired at the Cornell High Energy Synchrotron Source (CHESS) C1 beamline at 15.1 keV with a sample-detector distance of $\sim$ 1 m. The DNA/RNA and background (dialysis buffer) samples were held at 22 ± 1°C in a 3 mm path length cell with Si_3N_4 windows. For each sample, 8 exposures

of 10 s duration were recorded using a fiber-coupled CCD [105]and normalized by the incident intensity measured with a He ion chamber just before the sample. Images were azimuthally averaged and checked for time-dependent changes that might indicate radiation damage. Scattering angle was converted to momentum transfer ($q = 4\pi \sin(2\theta/2)/\lambda$) using the diffraction ring from a Silver stearate standard. For each image, the standard errors in intensity, $\Delta I(q_i)$, were estimated using the standard deviation σ_i among the N_i pixels in each q_i bin as $\Delta I(q_i) = \sigma_i/\sqrt{N_i}$ (for our detector geometry, $N_i \geq 100$ with 120 bins). Errors were propagated through subsequent averaging, buffer subtraction, and scaling operations.

Because single-stranded DNA is an extended, highly charged molecule, careful attention was paid to interparticle interference effects resulting from electrostatic repulsion. For data taken in each solution condition, SAXS curves at three DNA concentrations were matched at high q ($0.1 \leq q \leq 0.35\text{Å}^{-1}$) where interparticle interference effects are small, and the low q ($q < 0.1\text{Å}^{-1}$) region was compared. Interparticle interference was found to be negligible (within error) at a concentration of 75 μM for both RNA and DNA in 100 mM NaCl.

4.2.3 smFRET measurements

Since the first experimental realization of single-molecule fluorescence spectroscopy, the technique has gained immense popularity in the field of biophysics. Its main advantage is its ability to distinguish subpopulations while bulk measurements can only capture the ensemble average value. FRET occurs via dipole-dipole coupling between a donor and an acceptor fluorophore.

The efficiency of transfer depends on the sixth power of the separation of molecules, making it extremely sensitive to intra- and inter-molecular distances. Single molecule FRET (smFRET) then allows one to observe conformations of molecules individually. Experimental implementation of single molecule fluorescence measurements can be generally divided into two categories, where the molecules of interest are either freely diffusing through a fixed observation point (Figure 4.1 left) or spatially localized by a tether (Figure 4.1 right).

In the freely diffusing scheme, a very low concentration is used to ensure that at most one molecule passes through the focal volume and is detected at any time. Single molecule events then appear as spikes in the detected intensity trace as a function of time. In this scheme, the molecule can be observed unperturbed by tethers or surface interactions. However, this process may require long measurement times to collect enough events for a meaningful distribution. Furthermore, observation of each molecule is limited to its residence time within the focal volume; therefore, slower dynamics are unachievable this way.

In the spatially localized scheme, molecules are tethered to a surface via a chemical linker like a streptavidin-biotin pair. Using a wide-field imaging technique like total internal reflection fluorescence (TIRF) microscopy with a charge coupled diode (CCD) camera, single molecules appear as bright spots against a dark background. Thousands of molecules can then be observed simultaneously, thereby reducing measurement time. The observation time of a molecule is then limited only by the photobleaching time of the fluorophore, enabling even slow reaction dynamics on the order of seconds to be observed. However in some cases, the presence of a physical tether may interfere with the intrinsic property or behavior of the molecule of interest.

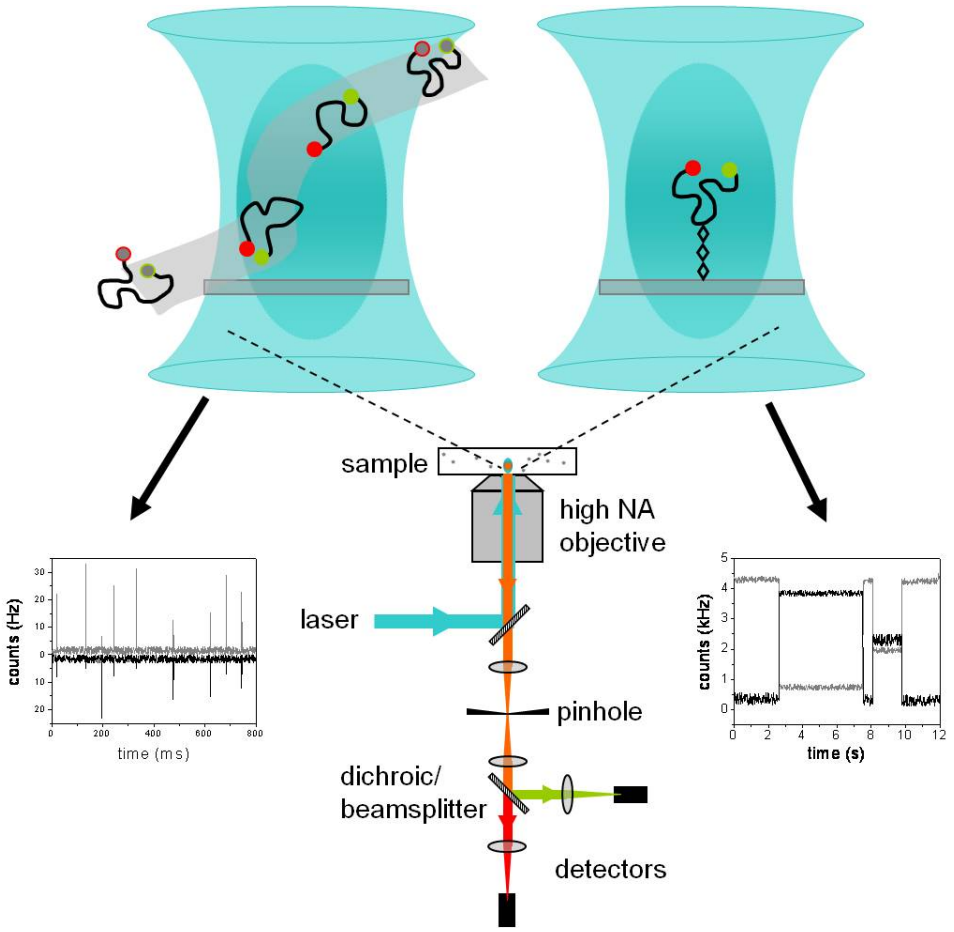

Figure 4.1: Schematic of typical smFRET setups. On the left, a stationary confocal observation volume probes molecules when they diffuse through. By working at extremely low concentrations, one can ensure that at most one molecule passes through the focal volume at any one time. In the other scheme, molecules are tethered to the surface and imaged simultaneously by wide field microscopy techniques like TIRF.

Our setup for smFRET measurements is shown on the left in Figure 4.1 . Briefly, a 532nm continuous wave laser is directed through a beam expander into an Olympus IX71 inverted microscope. A 60x UPlan-Apochromat objective focuses the beam down to a tight focal spot in the Nunc-tek sample chamber. The fluorescence is collected back from the same objective and separated from the excitation beam by a 550LP filter. The donor and acceptor fluorescence were further separated by a 660LP filter, and 570/40 and 670/40LP filters in the donor and acceptor channels respectively were put in to reduce cross talk and scatter of laser excitation light. Donor and acceptor fluorescence were separately collected through two 50 μm single-mode optical fibers from OZ Optics and detected by two single photon counting avalanche photodiodes from Perkin-Elmer Inc (AQ-SPCM). A Flex-2kD correlator card (Correlator.com, Bridgewater, NJ) in the Photon Counting mode samples the photon counts at 25 ns intervals and sends the data to a computer to be saved. Photon traces of each sample were recorded in 30 1-minute segments for a total of 30 minutes.

4.2.4 Determining single molecule events

Raw photon counts are processed and analyzed with a MATLAB routine attached in Appendix A. Briefly, photon counts are aggregated into 1 ms time bins with the size of the bins selected to be longer than the average residence time of a molecule in the focal volume ($\approx$ 500 μs). Individual molecules diffusing through the focal volume then appear as spikes in the intensity trace over time. The corrections described in the previous section were applied to the donor and acceptor intensities. Since double-labeling of molecules is complicated, the samples are always contaminated with molecules that have only the donor or the

acceptor fluorophore. Experimenter-selected threshold values applied to the intensity trace are often used not only to pick out the single molecule events, but also to separate the double-labeled molecules from the single-labeled ones [67, 109]. However, we found that not only do these subjective threshold values significantly change the calculated FRET efficiencies, they also have to be determined individually for each sample. Instead, we developed a criteria based on percentiles that we could apply consistently to all samples. First, we analyzed the donor, acceptor and total (donor plus acceptor) intensities separately, picking out the time bins with the top nth percentile values of each channel, where 'n' for each channel can be decided by the experimenter. We identified an event (ie., single molecule that is labeled with both donor and acceptor within the focal volume) when all three channels (donor, acceptor and sum) fulfilled their individual percentile criteria. Empirically we determined the percentile criteria as 95th, 99th and 95th in the donor, acceptor and sum channels respectively, and this criteria has worked well regardless of the samples or experimental conditions. The FRET efficiencies were calculated for each of these events and a population histogram was built. The distribution of FRET efficiencies was then fit with either one or more Gaussian distributions to get the average value.

4.2.5 smFRET calibration and corrections

To make accurate quantitative measurements of FRET efficiencies, several corrections have to be applied to the raw data:

1. Dark counts of the detector
2. Photons from background

3. Bleedthrough of fluorescence from the donor to the acceptor channels and vice versa
4. Difference in collection efficiencies in the donor and the acceptor channels
5. Difference in quantum efficiencies of the donor and acceptor fluorophores

Corrections for the dark counts of the detector (1) and background counts (2) can be made together by measuring the counts ($I_{backgroundA}$ and $I_{backgroundD}$) from a buffer-only sample and subtracting this from all subsequent measurements. Bleedthrough of fluorescence (3) is characterized by β, which is the fraction of photons that 'leaked' from the donor to the acceptor channels and vice versa. For example, the donor-to-acceptor bleedthrough is measured by dividing the photon counts in the acceptor channel by that in the donor channel when a donor fluorophore only sample is excited. Bleedthrough generally occurs more strongly in the donor-to-acceptor direction with minimal bleedthrough in the other direction; hence, bleedthrough into the donor channel will be assumed to be zero. Corrected intensities from the donor and acceptor channels (I'_D and I'_A respectively) can then be calculated from measured intensities (I_D and I_A) by:

$$I'_A = I_A - I_{backgroundA} - \beta(I_D - I_{backgroundD}) \tag{4.1}$$

$$I'_D = I_D - I_{backgroundD} + \beta(I_D - I_{backgroundD}) \tag{4.2}$$

Unfortunately, the differences in collection efficiencies (4) and quantum efficiencies (5) cannot easily be decoupled, but both corrections can be accomplished easily by one simple measurement. First, individual solutions of donor

and acceptor fluorophores need to be made up with concentrations such that at the laser excitation wavelength the two solutions have the same optical density, thus, the same absorption. Then, the photon counts collected from each solution is a multiplication of the collection efficiency (ϵ)of the system and the quantum efficiency (η) of the fluorophore. By dividing the counts from the donor fluorophore in the donor channel over that of the acceptor fluorophore in the acceptor channel, one is able to calculate γ, the relative difference in efficiencies.

$$\gamma = \frac{\eta_A \epsilon_A}{\eta_D \epsilon_D} \tag{4.3}$$

The corrected FRET efficiency is then given by:

$$E_{FRET} = \frac{I'_A}{I'_A + \gamma I'_D} \tag{4.4}$$

4.2.6 Measurement of Förster Radius

In order to make quantitative calculations of the separation between the 3′ and 5′ ends by FRET ($\langle R \rangle_{FRET}$), we have to accurately measure the Förster radius R_0 [102] of the Cy3-Cy5 pair. It is possible that the photophysical properties of Cy3 and Cy5 are different when they are attached to different nucleic acids (dexoribose vs. ribose), giving rise to different R_0 in ssDNA and ssRNA, and hence a different calculated $\langle R \rangle_{FRET}$. To verify that this was not the case, we measured R_0 using singly labeled dT_{10} and rU_{10}. The quantum yields of the Cy3-labeled nucleotides were measured against a Rhodamine 101 calibration standard. The quantum yields and the spectral overlap of the fluorophores were measured on

a spectrophotometer (CaryWinUV, Agilent Technologies, Santa Clara, CA) and a fluorimeter (PTI Inc, Birmingham, NJ). We observed that while there were no differences in the spectral overlap of the two fluorophores when labeled to different nucleic acids, Cy3-rU_{10} (QY = 0.181) had a lower quantum yield than Cy3-dT_{10} (QY = 0.205). As R_0 depends on $QY^{1/6}$, this resulted in a less than 3% difference in R_0 between ssDNA and ssRNA. We used our experimentally derived R_0 values of 56.4 Å for dT_{40} and 55.0 Å for rU_{40} in all our calculations of $\langle R \rangle_{FRET}$.

4.2.7 Polymer chain models

The worm-like chain (WLC) [50] is one of the most commonly used models to describe semi-flexible biological polymers. Unlike other models like the freely-jointed chain (FJC) or freely-rotating chain (FRC) models which involve rigid bonds, the WLC model describes a continuously flexible chain with a persistence length l_p over which correlations of direction of the tangent are lost. The probability distribution function of the end-to-end distance of a WLC has been derived [107]:

$$p(r,t) = \frac{4\pi A r^2}{(1-r^2)^{9/2}} \exp \frac{-3t}{4(1-r^2)} \tag{4.5}$$

The normalization constant A is given by:

$$A = \frac{4(3t/4)^{3/2} \exp 3t/4}{\pi^{3/2}(4 + 12/(3t/4) + 15/(3t/4)^2)} \tag{4.6}$$

where t=L/l_p, r=R/L, L is the contour length and l_p is the persistence length.

4.2.8 Data fitting

X-ray scattering curves were generated from the worm-like chain (WLC) model without excluded volume effects using a semi-analytical form factor, $I_{WLC}(q, L, l_p)$ [84]. To account for the finite thickness of the nucleic acid, the scattering function was modified by a cross-sectional scattering factor and is given by:

$$I(q) = I_0 \, I_{WLC}(q) \, \exp(-q^2 R_{cs}^2/2) \tag{4.7}$$

where R_{cs}^2 is the mean squared radius of gyration of the cross section.

WLC parameters L and l_p were converted into FRET efficiency using the theoretical end-to-end distance distribution function, $P_{WLC}(R, L, l_p)$,[107, 73]

$$E_{FRET} = \int_0^\infty dR \frac{1}{1 + (R/R_0)^6} P_{WLC}(R, L, l_p) \tag{4.8}$$

and the FRET-averaged end-to-end distance was calculated using:

$$\langle R \rangle_{FRET} = R_0 (1/E_{FRET} - 1)^{1/6} \tag{4.9}$$

In order to fit SAXS and smFRET data simultaneously to the WLC model, we minimized a combined χ^2 function consisting of the SAXS curve and FRET efficiency with weights given by our best estimate of experimental uncertainties. For SAXS, the $\Delta I(q_i)$ were propagated through the data reduction procedure, and for smFRET, an uncertainty of 2.2% was consistent with the standard deviation between independent measurements. The best-fit parameters l_p, R_{cs}, L, and

I_0 were found using nonlinear least squares (NLINFIT function in MATLAB). Confidence intervals were calculated from the estimated covariance matrix [87].

The radius of gyration is calculated from the fits using

$$\left\langle R_g^2 \right\rangle = \frac{L\, l_p}{3}\left[1 - 3\left(\frac{l_p}{L}\right) + 6\left(\frac{l_p}{L}\right)^2 - 6\left(\frac{l_p}{L}\right)^3\left(1 - \exp\left(\frac{-L}{l_p}\right)\right)\right] \tag{4.10}$$

4.3 Results

4.3.1 Compaction of dT_{40} and rU_{40} in $MgCl_2$ and NaCl revealed by smFRET

To compare the flexibilities of dT_{40} and rU_{40}, we measured the FRET-averaged end-to-end distance, $\langle R \rangle_{FRET}$, at varying concentrations of NaCl using smFRET. The efficiency of energy transfer (E_{FRET}) between Cy3 and Cy5 attached to the 3′ and 5′ ends of the nucleic acid chains was measured and converted to $\langle R \rangle_{FRET}$ using experimentally measured values of Forster radius R_0 (see Section 4.2.6 for details). FRET efficiency (E_{FRET}) histograms from smFRET of dT_{40} and rU_{40} at all conditions contained one peak (Figure 4.2), indicating that conformational fluctuations take place on time scales much faster than the dwell time of the molecule in the observation volume ($\sim 500\mu$s).

As we expected, $\langle R \rangle_{FRET}$ of both dT_{40} (circles) and rU_{40} (squares) decreased with increased salt concentration as excess cations reduced repulsion along the negatively charged backbone (Figure 4.3). $\langle R \rangle_{FRET}$ of both nucleic acids was indistinguishable below 400 mM NaCl, but above 400 mM NaCl, dT_{40} has a smaller $\langle R \rangle_{FRET}$ than rU_{40}. Assuming that the stiffness of a polyelectrolyte arises

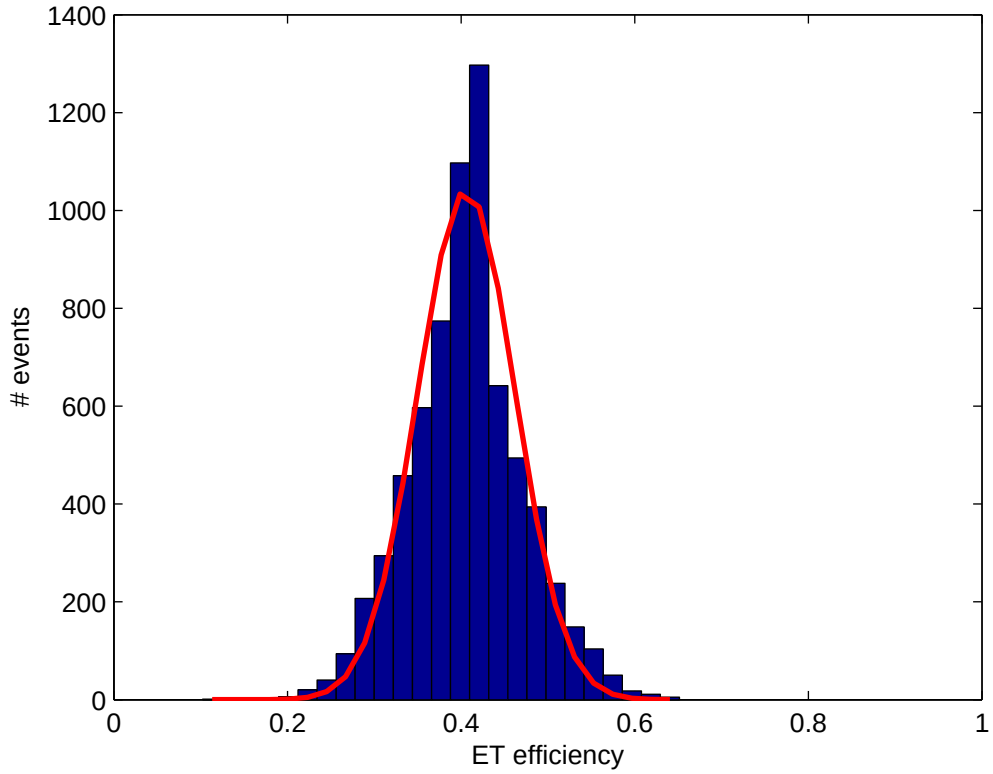

Figure 4.2: Typical population histogram (in blue) of smFRET measurements as shown by dT_{40} in 20mM Tris buffer at pH 8 with 10 mM $MgCl_2$. The histogram is fit to a Gaussian distribution (in red) and the position of the peak value is the average E_{FRET}.

only from the intrinsic flexibility of the backbone and repulsion due to electrostatics, then we expect the electrostatic contribution to be completely screened out at high ionic strengths. Thus, the observed divergence of end-to-end distance in high salt concentrations indicates differences in the intrinsic flexibilities of dT_{40} and rU_{40}.

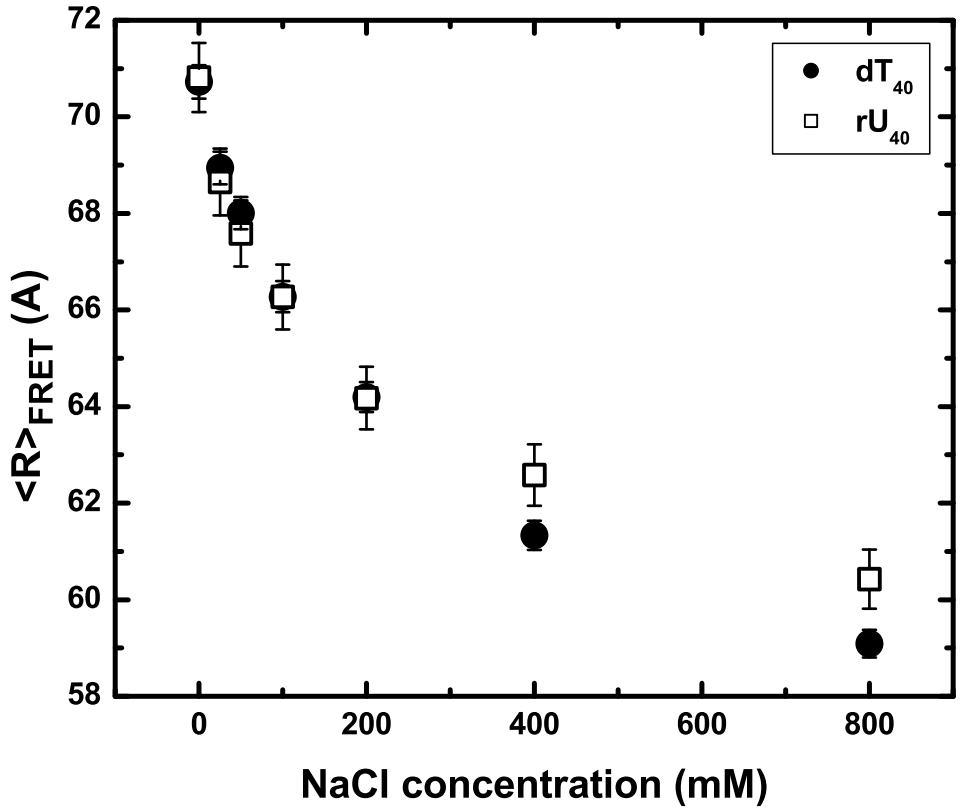

Figure 4.3: Dependence of the FRET-averaged end-to-end distance, $\langle R \rangle_{FRET} = R_0 (1/E_{FRET} - 1)^{1/6}$ $\langle R \rangle_{FRET}$ on the concentration of NaCl for dT_{40} (circles) and rU_{40} (squares). Error bars are standard errors from combined uncertainties in R_0 and E_{FRET}. The decrease in end-to-end distance is expected as excess cations screen the negatively charged nucleic acid backbone, resulting in greater chain flexibility at high salt. While rU_{40} and dT_{40}, are the same within error up to ~200 mM NaCl, they diverge at high salt, with dT_{40} appearing more compact.

4.3.2 Na^+ and Mg^{2+} have different screening efficiencies

To compare the charge screening efficiency of monovalent and divalent cations, we also measured $\langle R \rangle_{FRET}$ of dT_{40} and rU_{40} in the presence of increasing concentrations of $MgCl_2$, as shown in Figure 4.4. Similar to its behavior in NaCl, $\langle R \rangle_{FRET}$ of both dT_{40} and rU_{40} decreased as the concentration of $MgCl_2$ increased. At the same ionic strength, both dT_{40} and rU_{40} had a smaller$\langle R \rangle_{FRET}$ in Mg^{2+} than in Na^+, indicating that the nucleic acids were more collapsed in Mg^{2+} than in Na^+.

4.3.3 Local environment can perturb the flexibility of nucleic acids

We investigated the effect that the local environment has on the flexibility of the single-stranded nucleic acid by measuring $\langle R \rangle_{FRET}$ of dT_{40} with an 18-mer double-stranded helix attached to one end ($dT_{40} - duplex$, see inset of Figure 4.5) at different NaCl concentrations. Compared to the single stranded dT_{40} construct, the presence of a helix at one end of a dT_{40} chain results in a lower E_{FRET} or a larger $\langle R \rangle_{FRET}$. This indicates that the conformation of the nucleic acid chain is very sensitive to the local environment and the presence of a tethering helix can cause the dT_{40} chain to appear more extended.

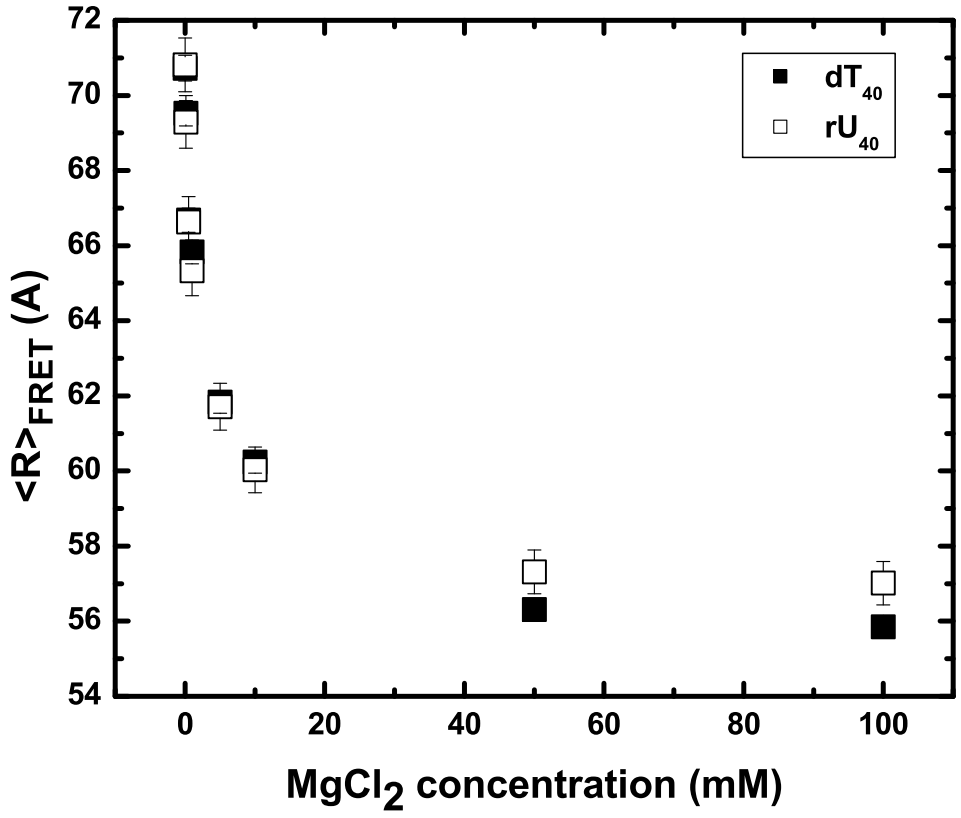

Figure 4.4: Dependence of $\langle R \rangle_{FRET}$ on the concentration of $MgCl_2$ for dT_{40} (filled squares) and rU_{40} (unfilled squares).

4.3.4 Determination of WLC model parameters for dT_{40} and rU_{40} using smFRET and SAXS data

While it is clear from smFRET measurements that dT_{40} and rU_{40} have different properties, and that Mg^{2+} is more effective at charge screening than Na^{+} at the same ionic strength, further interpretation of the FRET data requires a model for

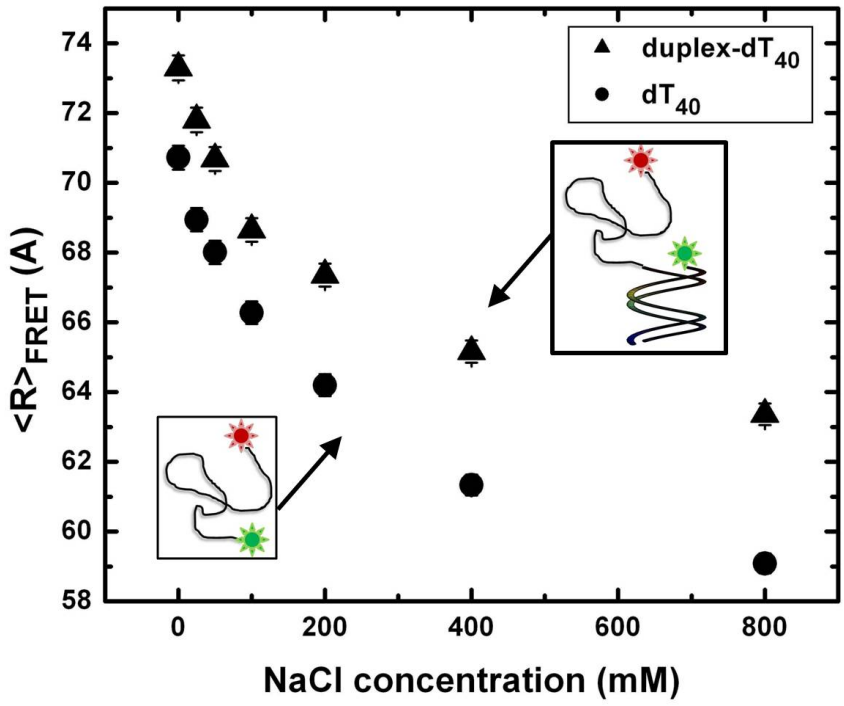

Figure 4.5: Effect of a duplex attached to the single strand. The presence of an attached duplex changes the measured flexibility of ss-DNA. The 40-mer construct with (dT_{40} – *duplex*, triangles) and without (dT_{40}, circles) an attached 18bp duplex DNA, exhibit different smFRET-averaged end-to-end distances, $\langle R \rangle_{FRET}$ at all NaCl concentrations measured. This indicates that the duplex excludes the single stranded chain from its vicinity, and may explain the discrepancy between our smFRET measurements of freely diffusing dT_{40} and previous studies using tethered DNA that report more extended chains.

the end-to-end distance distribution. The WLC model uses two parameters to specify this distribution: (1) the contour length, L, and (2) the persistence length l_p. The fact that smFRET measures only one parameter (the FRET efficiency) means that additional information is required to determine the polymer properties. For the same reason, the FRET efficiency cannot be used to determine whether WLC is an appropriate model. To address both concerns, we additionally performed SAXS measurements of dT_{40} and rU_{40} at 100 mM NaCl and asked whether the WLC model could account for the entire scattering curve in addition to the smFRET data.

The SAXS and smFRET data at 100mM NaCl were fit simultaneously to the WLC model for each nucleic acid using nonlinear optimization, where semi-analytical methods were used to calculate the FRET efficiency [73] and scattering curve [84] from model parameters (see Methods). The WLC scattering curves are in excellent agreement with the SAXS data while predicting the FRET data within the experimental uncertainty (Figure 4.6). The values of L and l_p that allowed the WLC model to fit the experimental data to at least 95% confidence are shown by the shaded ellipses in the inset of Figure 4.6. These acceptable values for dT_{40} and rU_{40} are represented by non-overlapping regions in L-l_p space, confirming that rU and dT have significantly different flexibilities despite having similar $\langle R \rangle_{FRET}$ at low ionic strength.

A cartoon drawing of ssDNA and ssRNA illustrating that RNA has a shorter interbase distance and is less flexible is shown in Figure 4.7.

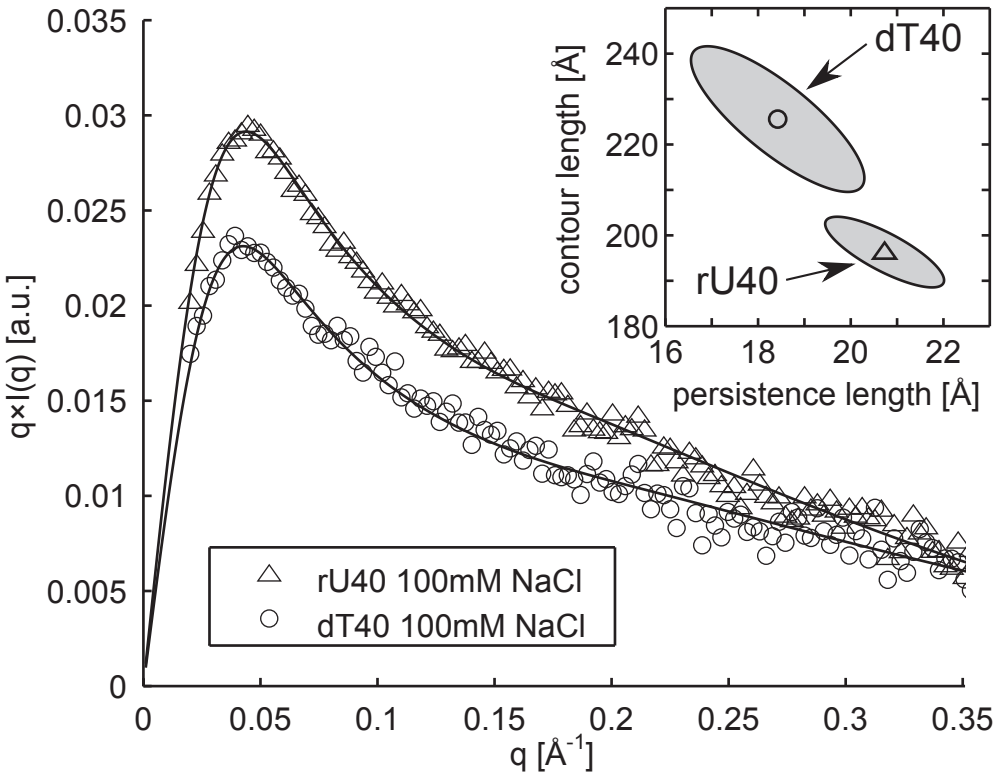

Figure 4.6: SAXS curves for 75μM dT_{40} (circles) and rU_{40} (triangles) in 100mM NaCl are shown together with a worm-like chain (WLC) model for the scattering (solid lines) that optimizes the fit to SAXS data and smFRET data ($\langle R \rangle_{FRET}$ = 66.3±1.5Åfor both rU and dT). Experimental uncertainties were used as weighs in the fit. The best-fit values of L and l_p predict $\langle R \rangle_{FRET}$ values of 68.0Å for dT and 67.2Å for rU. The WLC scattering curves include an overal scale factor and an additional free parameter, R_{cs}, to correct for finite molecular cross-section (see Section 4.2.8) equal to 4.2Å for rU and 3.7Åfor dT. In the inset, the persistence length l_p and the contour length L of each wormlike chain are shown as symbols with ellipses that enclose the 95% confidence interval. The non-overlapping ellipses show that dT_{40} and rU_{40} are described by wormlike chains with significantly different properties: dT_{40} has a longer L and a shorter l_p than rU_{40}.

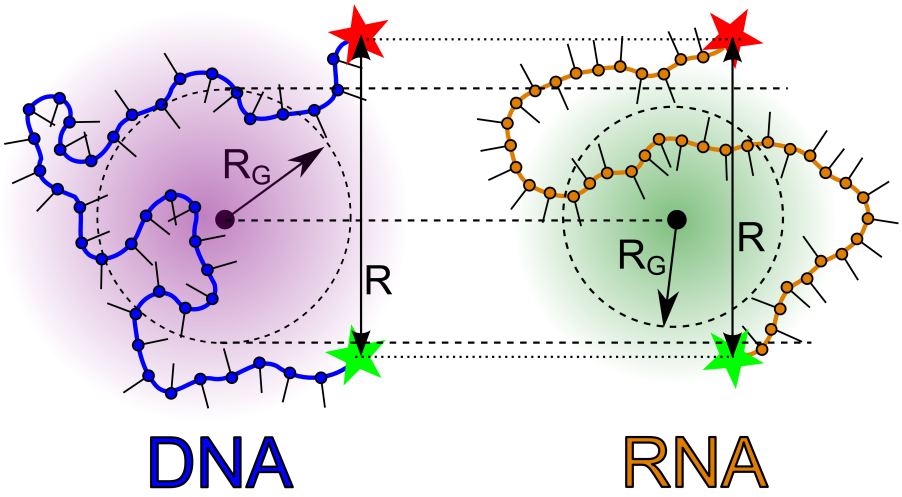

Figure 4.7: Cartoon drawing of ssDNA and ssRNA with different polymer properties. While ssDNA has a longer contour length L, it has a shorter persistence length l_p.

4.3.5 The ion-dependent flexibilities of rU and dT

The WLC model with best-fit contour lengths, $L = 196.4$ and $L = 225.6$Å for rU_{40} and dT_{40} respectively, was used to convert $\langle R \rangle_{FRET}$ (Figures 4.4 and 4.3) to l_p over a wide range of ionic conditions in Figure 4.8. The ionic strength (I) dependence of l_p of each nucleic acid is well described by a phemenologically-derived function,

$$l_p(I) = l_p(\infty) + \frac{l_p(0) - l_p(\infty)}{(b\,I)^{n/2} + 1} \tag{4.11}$$

The equation satisfies the requirements of (1) a power-law limiting behavior at high ionic strength ($I > b^{-1}$) of $l_p(I) - l_p(\infty) \sim I^{-n/2}$, and (2) a finite persistence length in no excess salt. The parameter, b, represents the charge screening efficiency and is equal to the inverse of the ionic strength at the midpoint between the most extended and compact states of the molecule. The exponent, n, is an attempt to capture the power law dependence of the persistence length on I, often predicted from polyelectrolyte theories where the electrostatic contribution to persistence length l_e depends on the Debye screening length $\kappa \sim I^{1/2}$ as $l_e \sim \kappa^{-n}$. We fit our measured values of $\langle R \rangle_{FRET}$ to Equation 4.11 to obtain b, n, $l_p(\infty)$ and $l_p(0)$ in both Na^+ and Mg^{2+}. We also imposed the condition that $l_p(\infty)$ and $l_p(0)$ values for rU_{40} or dT_{40} had to be the same in both salts, but could vary between the nucleic acids. These results are shown in Figure 4.8.

4.4 Discussion

4.4.1 Understanding the differences between ssDNA and ssRNA

Using two complementary methods, SAXS and smFRET, we observe that dT_{40} and rU_{40} have significantly different conformations in solution. This is evident from the smFRET data at high ionic strengths, where electrostatic repulsion is expected to be mostly screened by the ions and thus $\langle R \rangle_{FRET}$ reflects the intrinsic backbone flexibility. This difference between rU and dT can also be seen when SAXS and smFRET data for the two nucleic acids in 100 mM NaCl are used to constrain the worm-like chain model. In this case, RNA and DNA are best de-

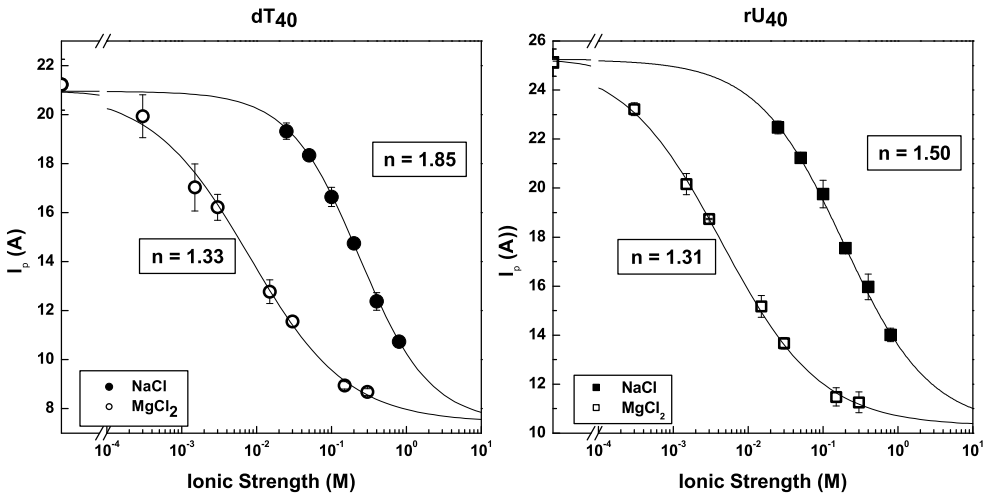

Figure 4.8: The ionic strength dependence of the persistence length (l_p) for dT_{40} (left panel) and rU_{40} (right panel) is different in the presence of monovalent NaCl (filled symbols) and divalent $MgCl_2$ (unfilled symbols). The ionic strength is represented on a logarithmic scale. Solid lines indicate fits to a phenomenological model for the ionic strength dependence (see Equation 4.11). For dT_{40}: $b_{Na^+} = 4.25\,M^{-1}$, $b_{Mg^{2+}} = 124\,M^{-1}$, $l_p(\infty) = 7.44$, $l_p(0) = 20.9$. For rU_{40}: $b_{Na^+} = 5.35\,M^{-1}$, $b_{Mg^{2+}} = 224\,M^{-1}$, $l_p(\infty) = 10.3$, $l_p(0) = 25.2$. The scaling exponent, n, is shown in the figure.

scribed by two significantly different worm-like chains, as illustrated in Figure 4.6.

Our finding that rU_{40} has a shorter L and longer l_p is consistent with how chemical differences between the nucleic acids are known to affect backbone conformations. Sugar pucker can determine the interphosphate distance between adjacent bases, with C'3 endo producing a more compressed backbone than C'2 endo [28]. NMR studies have shown that the sugar pucker in poly(dT)

is predominantly in the C'2 endo state [11] while poly(rU) is a 44:56 mixture of C'2 endo and C'3 endo [57]. Thus, we expect rU_{40} to have a shorter contour length. Also, while rU_{40} is known to undergo the least stacking amongst the ribonucleotides, there is some evidence that rU_{40} does in fact take part in some base stacking [118, 57]. We believe that base stacking in poly(rU) but not in poly(dT) may cause rU_{40} to be more rigid than dT_{40}. Ultimately, both sugar pucker and base stacking effects have the effect of making rU_{40} shorter and stiffer than dT_{40}.

4.4.2 A molecular tether can perturb smFRET measurements

Published literature on the flexibilities of ssDNA and ssRNA reveals a rather large range of persistence lengths that span 10-60 Å, measured using different techniques under a variety of conditions [16, 108, 56, 71]. While our measurements of l_p (11-22Å for dT_{40}) fall within this range, they are on the low end even when compared to similar studies using smFRET. Perhaps most meaningful is a comparison to Reference [73], a work that not only measured the E_{FRET} of dT_{40} tethered to a glass surface at the same ionic strengths using smFRET, but also used a similar WLC analysis. Interestingly, the authors found that l_p was 15-30 Å between 0 and 2 M NaCl, which is consistently ~ 50% higher than our values over the entire range of ionic strength. Our measurements of $dT_{40} - duplex$ show why this difference might arise (Figure 4.5). The presence of a highly charged duplex nearby, like the double-stranded tether used to immobilize the ssDNA, can affect the $\langle R \rangle_{FRET}$ measured. The strong electrostatic repulsion from the duplex could reduce the space available to the single strand by electrostatic excluded volume [1], or by stiffening the chain in the vicinity

of the duplex. Our measurements of dT_{40}, unencumbered by the repulsion of the duplex, find a shorter l_p, a lower $\langle R \rangle_{FRET}$, and a more compact chain. Interestingly, our values for $dT_{40} - duplex$ do not completely recover the values observed in Reference[73], indicating that surface effects may also play a role. Our results indicate that the effect of the tether, and possibly that of the glass surface nearby, can have a strong effect on chain properties, and highlights the extra care that experimenters should take when interpreting results from tethered molecules. The observation that polymer conformations can be perturbed by adjoining duplexes also has implications for our understanding of biologically relevant nucleic acids, which are often made up of interspersed regions of unpaired and paired bases.

4.4.3 Beyond the Debye-Hückel theory

The differences in chain conformations in the presence of Na^+ or Mg^{2+} can be explained in terms of nucleic acid-ion interactions. For both RNA and DNA, we find that Mg^{2+} is ~20-40 times more efficient at charge screening than Na^+ in terms of ionic strength (60-120 times in terms of cation concentration). This relative screening efficiency agrees well with the ~17-83 times observed in single molecule force spectroscopy of chemically denatured ssDNA [68], although it disagrees with predictions from Debye-Hückel (DH) theory that screening efficiency should be determined solely by ionic strength. We have also previously measured the repulsion between 25bp duplex DNA in Na^+ and Mg^{2+}, and found that 3 mM $MgCl_2$ had the equivalent screening effect to 150 mM NaCl [90], which yields an ionic strength ratio of ~ 17.

Polyelectrolyte theories for charged worm-like chains in salt solutions predict a general power-law dependence of the electrostatic component of the persistence length on the Debye screening factor. The value of the scaling exponent, n, varies from 1 to 2 depending on the assumptions of the model. To compare our results with predictions from polyelectrolyte theory, we fit our data using Equation 4.11, a formula designed to capture the asymptotic power-law dependence at high ionic strength where these theories are valid. Theoretical work from Ha and Thirumalai [30] shows unambiguously which scaling theories apply depending on the relative length scales in the problem (l_0 and the charge spacing, d). Assuming $l_0 = l_p(\infty)$ and $d = L/40$, ssDNA and ssRNA should behave like flexible polyelectrolytes with electrostatic persistence lengths given by the Odijk-Skolnic-Fixman (OSF) [99, 78] theory above ~60 mM ionic strength. Indeed, we observe a crossover to power-law scaling around this point for NaCl (i.e. $1/d$ from fits to Equation 4.11). However, from the OSF theory, we expect a power law exponent of $n = 2$ in Equation 4.11 that does not agree with the data: for rU and dT in NaCl, we observe $n = 1.50$ and 1.85 respectively, both lower than predicted.

The non-OSF scaling may indicate that more complex ion-nucleic acid interactions are important for charge screening than are captured by the Debye-Hückel (DH) approximation from which OSF theory is derived. Netz and Orland [76] go beyond DH theory and show that salt-dependent charge renormalization can lead to an apparent scaling factor $n < 2$ that becomes less ideal (smaller) for chains with higher linear charge density. Intriguingly, we observe a smaller value of n for rU than for dT, a trend that is consistent with the charge renormalization theory given our measurement of a shorter phosphate distance in rU. However, this finding remains tentative because apparent $n < 2$ scaling

can also be caused by excluded volume effects and finite chain length [64]. We believe this interesting result merits further study.

4.4.4 Suitability of the WLC model

Throughout this study, we relied on the WLC model to parameterize the distribution of conformations sampled by the nucleic acid in equilibrium. This is partly justified by the excellent agreement between WLC-derived scattering curves and SAXS data, as well as $\langle R \rangle_{FRET}$ predictions, for both rU_{40} and dT_{40}. Furthermore, electrostatic theories based on WLC are tractable for experimentalists, making this a popular model for nucleic acids and proteins. However, WLC is not general enough to describe ssDNA or ssRNA under all conditions. We anticipate that all-atom or coarse grained methods combined with realistic treatment of electrostatics will lead to improved modeling of flexible RNA. Complementary techniques including SAXS and smFRET will be invaluable in this effort to build better models going forward.

4.5 Conclusion

A combination of SAXS and smFRET revealed conformational differences between model single-stranded nucleic acids lacking secondary structure (dT_{40} and rU_{40}). In 100mM NaCl, we showed that while rU_{40} and dT_{40} resemble worm-like chains, rU_{40} has shorter contour length (L) and longer persistence length (l_p) than dT_{40}. Applying this model to smFRET data, we find that rU_{40} has longer l_p than dT_{40} in Mg^{2+} and Na^+ at all ionic strengths measured. Fur-

thermore, for both nucleic acids, we find that one needs 20-40 times lower ionic strength of $MgCl_2$ than NaCl to achieve the same chain compaction. Similarly, we probed the effect of the local environment of chain flexibility, showing that the presence of a flanking double-stranded helix affects the conformation of the single-stranded region. This has implications for biologically-relevant nucleic acids with interspersed regions of paired and unpaired bases. Future directions should extend this work by introducing different bases or mixed sequences to understand the effect of base stacking on chain flexibility. The difference in l_p between ssRNA and ssDNA and the dramatic ion dependence of conformation are likely to affect folding and interactions with binding partners in the cell.

CHAPTER 5

CONCLUSION

Throughout the past six years of research, I have not always been faithful in pursuing the original goals laid out at the beginning. The nature of original research is such that new opportunities and directions will appear from time to time, leading the open-minded and intrepid researcher down exciting new paths. I have enjoyed immensely those unexpected turns, and this dissertation is a reflection of such a journey. In this chapter I will describe recent developments in the field that make the results described in this dissertation significant, and speculate on further directions that could be explored.

The work described in Chapter 2 was motivated by amyloid structures, or protein aggregates that have been implicated in a growing list of human diseases like Alzheimer's, Parkinson's and Huntington's diseases [12]. Alzheimer's disease alone is projected to afflict 1 in 85 persons worldwide by 2050 [5]. In these diseases, protein aggregates are observed that consist of parallel or anti-parallel β-sheet structures that stack to form fibrillar aggregates with the β-sheet structures perpendicular to the fibril axis [103]. These "cross-β" structures were first predicted by Pauling and Corey in 1951 [83].

The constituent proteins that aggregate to form amyloid differ from disease to disease, yet despite the lack of sequence homology between these proteins, they are able to form the same cross-β structures. This observation led scientists to believe that just like α-helices or β-sheets, the cross-β amyloid structure is universal amongst proteins. This hypothesis was furthered by a classic experiment that showed that under the right conditions apomyoglobin (apoMb), which is neither known to form amyloid structures in nature nor associated with any dis-

eases, can form amyloid structures [25]. Furthermore, the role of the amyloid structures in diseases (cause or effect) is not known. They were once thought to be the cause of neuronal loss in neurodegenerative diseases. More recently, toxicity in cells have been attributed to soluble oligomers rather than large aggregates, leading to the hypothesis that amyloid aggregates may actually be beneficial by sequestering the toxic soluble species [6, 32].

The results I described in Chapter 2 do not relate directly to amyloid formation, but they provide a proof of principle that one can measure the conformational fluctuations in apoMb using FCS. The next logical step will be to perform the same measurements under the conditions where apoMb forms amyloid. The mechanisms by which proteins transition from their native states to the amyloid structures, and the triggers of this process are not well understood. I hope that by measuring the changes in fluctuations during the transition, we can isolate fluctuations that are directly involved in the transition.

While the follow-up experiments seem simple, I anticipate great difficulties in the analysis of the FCS data. FCS theory for equilibrium systems have been well described and the equations for the exepcted correlation functions of several phenomena such as anomalous diffusion and flow are straightforward. However, correlation functions for non-equilibrium systems or multivariable systems are neither trivial nor well described in literature [89]. Furthermore, a biological system like aggregation is unlikely to give clean data with well defined beginning and end states, making data analysis even more complicated. My current research will attempt to tweeze out some of these answers.

To extend FCS into a general technique capable of probing intramolecular fluctuations in other proteins, we need to know not just the identity of the

quencher, but also the mechanism of quenching. Altered photophysical properties of the fluorophore after covalent labeling is often an afterthought in experimental design, yet it has to be accounted for if quantitative measurements of fluorescence are desired. Thus, I set out to characterize the quenching of Alexa488 by the natural amino acids as detailed in Chapter 3. The surprising result that up to 4 amino acids can quenching Alexa488 led me to extend this study to other fluorophores. I believed that a compendium of such effects for all the commonly used fluorophores will be a useful guide for anyone using fluorescence techniques.

The creation of a compendium was going to be a repetitive task, and I was fortunate to have the help of a talented undergraduate. However, the lack of time, the lure of other more interesting projects and the repetitiveness from the sheer number of dyes that I had hoped to screen eventually put a damper on my plans. Instead of pursuing breadth, I changed the focus to really understanding the mechanisms of quenching, which had not been well documented. It is still my hope that someday someone will complete the characterization of quenching for many more dyes. This information is useful not only as a consideration when interpreting quantitative measurements of fluorescence, but also as a possible probe of molecular contacts and distances. It will also guide researchers in selecting labeling sites within a protein to.

The development of superresolution techniques like Photoactivated Localizataion Microscopy (PALM) [35, 3], Stimulated Emission Depletion (STED) microscopy [34] and Stochastic Optical Reconstruction Microscopy (STORM) [95] in the past few years has reenergized the field, allowing researcher to see smaller cellular structures with finer detail. PALM and STORM are based on the same

fundamental principle, that by switching fluorophores on and off at will, one can localize the position of a small subset of fluorophores within a sample with higher accuracy. By iterating the process, one can build up a entire image using a pointillism approach. The promise of these techniques has led to resurgent interest in creating new fluorophores that can be switched on and off [2], and our finding that the popular Alexa family of dyes can be quenched by specific amino acids upon contact is relevant.

The work in Chapter 4 was borne out of a long-standing collaboration with Professor Lois Pollack. During the search for conformational dynamics in functional nucleic acids like pseudoknots and ribozymes, we often faced with a fundamental question - what is the intrinsic flexibility of RNA? The answer to this question could not easily be found in published literature. Our study is timely because recent discoveries of functional RNAs that modulate essential biological processes like transcription and translation. The growing importance of RNAs is best illustrated by the number of Nobel Prizes awarded in recent years to RNA-related research like RNA interference (RNAi) [26]. For example, microRNA are short RNA strands that bind to mRNA transcripts and inhibit translation in eukaryotes [48] while small nuclear RNA (snRNA) are involved in RNA splicing. The breakthroughs so far in the RNA world have all been related to the discovery of new RNA molecules and biological functions. However, the biophysical properties of the RNAs have not been well studied. Knowledge of the fundamental properties of RNAs will become important when we start to probe the mechanisms by which these molecules perform their function.

Studies on the mechanical properties of DNA had been performed using single molecule fluorescence [73, 56] and force spectroscopy [101, 100], and

we used these results as a comparison when we extended the studies to RNA. We found differences between our results and those from published literature, which resulted from the experimental technique used. Specifically, molecular tethers used to spatially localize the target molecular for single molecule visualization were perturbing the target molecule. This finding is significant as the majority of single molecule fluorescence studies employ schemes that require tethering of the target molecules, so as to lengthen the observation time. While it is useful for most purposes, our study highlights the need to ensure that the tether does not perturb the system of interest. There has been some efforts to design alternative schemes to this end. Amongst them, Anti-Brownian Electrophoretic (ABEL) trap is one such development that uses tiny electric fields to trap single molecules in a small region by detecting and countering its Brownian motion in real time [13], eliminating the need for a physical tether. However, this technique is relatively new and has not yet achieved widespread use. With a reliable way to spatially localize a target molecule, we can potentially measure the properties of ssRNA more accurately by elucidating the full conformational probability distribution function.

Our findings detailed in Chapter 4, lays the groundwork for further experiments by detailing the pitfalls encountered in some single molecule methodologies. So far we have only measured the intrinsic flexibility of the chains using the poly-rU and poly-dT as our simplest models of chains with no base stacking. In follow-up experiments, it will be interesting to introduce chains with different bases to investigate the effect of base stacking on chain flexibility. One can also use mixed sequences to probe the dependence of chain flexibility on base sequence. Further experiments using different lengths of nucleic acid chains can also help to identify the most appropriate model (worm-like chain or freely

jointed chain etc.) for fitting ssRNA.

APPENDIX A

DIFFERENTIATING BETWEEN THE 1D+2F AND 1D+3F MODELS

To determine the model suitable for fitting FCS data at each pH across the entire series, the χ^2 values of the 1D+3F fit were subtracted from the χ^2 values of the 1D+2F fit for each pH, as depicted in Figure A.1. Above pH 4.35, the difference in χ^2 values is negligible, indicating that the 1D+3F model is not an improvement over the 1D+2F model. However, below pH 4.1, the difference in χ^2 values jumped abruptly, indicating that the 1D+3F model improved the fit. This justifies our use of two different fitting models over the entire pH series.

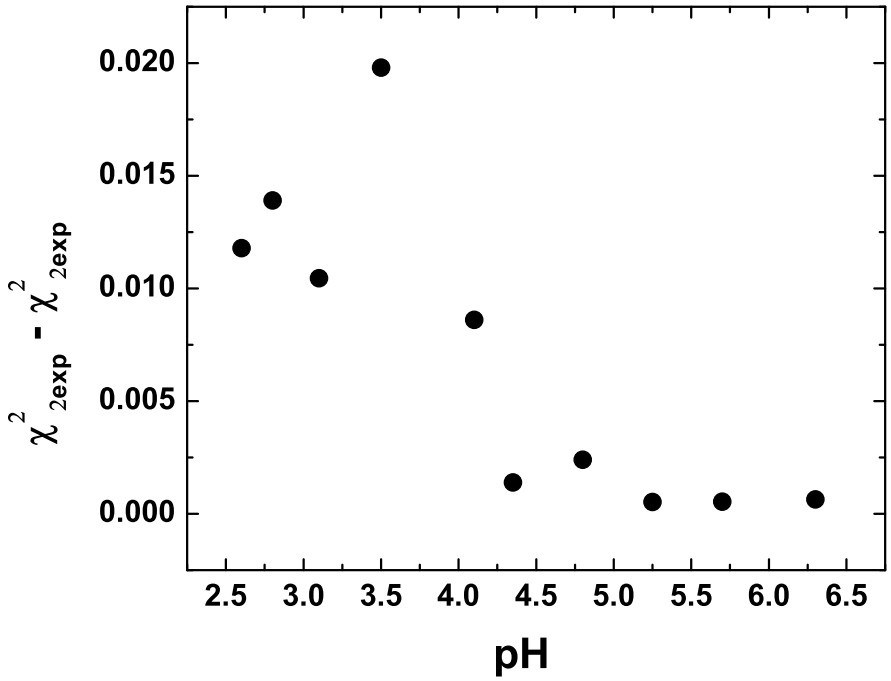

Figure A.1: Differentiating between the 1D+2F and 1D+3F models. The difference in χ^2 values between the two models were calculated across the entire pH series. A distinct jump at pH 4.1 justifies the use of two different fitting models across the pH series.

APPENDIX B

ALGORITHM FOR ANALYZING SMFRET TRACES

The steps in this MATLAB algorithm were briefly outlined in Chapter 4. The main program (etEff.m) in turn calls on three subroutines (readphotonstream.m, convertphotonstream.m and binphotons.m).

B.1 etEff.m

```
% This program takes the two .dat files 'acceptor' and
% 'donor' to create a histogram of the FRET efficiency.
% The program calls subroutines readphotonstream.m,
% convertphotonstream.m, and binphotons.m file.

% Enter the name of the photon stream files
nameString = 'dT40_Na100mM1_';

% Define the correction factors beta and gamma
gamma = 1.27;
beta = 0.067;

% Define the background counts in the donor and acceptor
% channels
backAcc = 0.198;
backDon = 0.16;

% Define the percentile levels in the donor, acceptor and
```

```
% total that will be used to determine single molecule
% events
Donperctile = 95;
Accperctile = 98;
Totperctile = 95;

for i = 1:30

% reads in data from user-specified file, converts
% photon countsinto intensity vs time
    string = [nameString, int2str(i), 'A.dat'];
    stream = readphotonstream(string);
    photontrace = convertphotonstream(stream);
    Astruc = binphotons(photontrace);
    intDonor1 = Astruc.trace;       % channel A=donor

    string = [nameString, int2str(i), 'B.dat'];
    stream = readphotonstream(string);
    photontrace = convertphotonstream(stream);
    Dstruc = binphotons(photontrace);
    intAcceptor1 = Dstruc.trace;    % channel B=acceptor

% subtract background and correct for beta and gamma
    intAcceptor1=(intAcceptor1-backAcc)-beta.*
    (intDonor1-backDon);
    intDonor1=(1+beta).*(intDonor1-backDon);
```

```
    intTot = intAcceptor1 + intDonor1;

% Pick out bins with the top percentiles as single
% molecule events
    Accpercentile = prctile(intAcceptor1,Accperctile);
    Donpercentile = prctile(intDonor1,Donperctile);
    Totpercentile = prctile(intTot, Totperctile);

% Find events that span multiple time bins and count
% those as only one event
    c1 = (intAcceptor1 > Accpercentile);
    c2 = (intDonor1 > Donpercentile);
    c3 = (intTot > Totpercentile);

    c2 = c1&c2;
    c = c2&c3;

    sum(c)

  % We will add a '0' to c on both sides, then add the
% two. Find [1 2 1]s. Then Use c==2 to get new c then
%  take out zero on either side. then add the two. Find
% the ones and twos. Remove the ones on the initial
% conditions and then keep the twos. Add in the [1 2 1]s.
% [1 2 1] takes care of even number of 1s whil ec==2
% condition takes care of odd and shortens the width of
```

```
% thespikes. Repeat until no more twos.
    sum1 = 1;
    while (sum1 > 0)
                          % c = [1 0 0 1 1 1 0 1 1]
     c1 = [0 c'];          % = [0 1 0 0 1 1 1 0 1 1]
     c2 = [c' 0];          % = [1 0 0 1 1 1 0 1 1 0]
     cond = c1 + c2;       % = [1 1 0 1 2 2 1 1 2 1]
     length = size(c);
     l = length(1);
     condeven1 = [0 cond 0];
     condeven2 = [cond 0 0];
     condeven3 = [cond(2:l+1) 0 0 0];
     ceven=((condeven1 ==1)&(condeven2 ==2)&
(condeven3 == 1));
     ceven = ceven(1:l);
     cond2 = (cond == 2);
     sum1 = sum(cond2);
     cond3 = cond2(2:l+1);
     cond4 = cond2(1:l);
     condodd = cond3 +cond4;
     c(find(condodd ==1)) = 0;
     c(find(ceven ==1)) = 1;
    end

    clear c1 c2 cond cond2 cond3 cond4 condodd condeven1
    condeven2 condeven3 ceven;
```

```
    sum(c)

    intAcc = intAcceptor1(c);
    intDon = intDonor1(c);

    intAcc = nonzeros(intAcc);
    intDon = nonzeros(intDon);

    clear c intAcceptor1 intDonor1;

    intDonor = [intDonor; intDon];
    intAcceptor = [intAcceptor; intAcc];

  clear intDon intAcc;

end

size(intAcceptor);
size(intDonor);
Et = intAcceptor./(intAcceptor + gamma*intDonor);


% Plot histogram
figure
 hist(Et,25)
  [N,x] = hist(Et,25);
```

```
 xlim([0 1]);
 hold on
t = [N;x];
fid = fopen('gaussfit.dat','w');
fprintf(fid, '%4i   %8.4f\n', t);
fclose(fid);

% Fit with one gaussian
x0=fminsearch('gafit',[1000 5 0.3]);
hold on;
ygauss = x0(1)*exp(-(x0(2)*(x-x0(3))).^2);
fitplot = plot(x, ygauss, 'r')
hold off;
xlabel('ET efficiency');
ylabel('# events');
set(fitplot, 'Linewidth', 2);
 x0
```

B.2 readphotonstream.m

```
function filename = readphotonstream(name)

% This function reads data file from correlator.com
% Photon.exe. The data is 8-bit. The first byte codes
% data. The second byte for either 8 or 16 bit states
% the system clock, either 40 or 60MHz.
```

```
% output.bitbyte = 8 or 16 bit data
% output.systemclock = 40 or 60MHz
% output.data = clockticks between photon detection events

fid = fopen(name);
headers = fread(fid, 2);
filename.bitbyte = headers(1);
filename.systemclock = headers(2);

filename.data = fread(fid);
fclose(fid);
```

B.3 convertphotonstream.m

```
function allphotons = convertphotonstream(filename)

% This function takes data file extracted by
% readphotonstream.m and converts it from clocktick
% of photon event to time of photon event.
% output = array of timepoints when each photon was detected

timebins=1./(filename.systemclock.*1E6);

% find index of array where there is photon
index=filename.data~=255;
```

```
%cummulative sum of clockticks
cummindex = cumsum(filename.data);

% find clockticks where there is a photon event
cummtime = index .*cummindex;

% convert clockticks to time traces
converted = cummtime .* timebins;

allphotons = nonzeros(converted);
```

B.4 binphotons.m

```
function intensity = binphotons(photontrace)

% This function takes data output from
% convertphotonstream.m and bins the photon events
% into regular timebins. User is allowed to select
% the sizeof the timebin. A histogram plot of the
% events is also given.
% output.bins = time (x-axis)
% output.trace = number of photons within said timebin

choice =9;

% choice = input('Please select binsize: 1. 0.1us;
```

```
% 2. 0.5us; 3. 1us; 4. 5us; 5. 10us; 6. 50us;
% 7. 100us; 8. 500us: ');

if choice == 1
    binsize = 1E-7;
elseif choice == 2
    binsize = 5E-7;
elseif choice == 3
    binsize = 1E-6;
elseif choice == 4
    binsize = 5E-6;
elseif choice == 5
    binsize = 1E-5;
elseif choice == 6
    binsize = 5E-5;
elseif choice == 7
    binsize = 1E-4;
elseif choice == 8
    binsize = 5E-4;
elseif choice == 9
    binsize = 1E-3;
elseif choice == 10
    binsize = 2E-3;
elseif choice == 11
    binsize = 1E-2;
elseif choice == 12
```

```
    binsize = 5E-2;
elseif choice == 13
    binsize = 1;
end

intensity.bins = 0:binsize:60;
intensity.trace = histc(photontrace, intensity.bins);
```

BIBLIOGRAPHY

[1] Y Bai, VB Chu, J Lipfert, VS Pande, D Herschlag, and S Doniach. Critical assessment of nucleic acid electrostatics via experimental and computational investigation of an unfolded state ensemble. *Journal of the American Chemical Society*, 130(37):12334–12341, 2008.

[2] M Bates, B Huang, GT Dempsey, and X Zhuang. Multicolor super-resolution imaging with photo-switchable fluorescent probes. *Science*, 317(5845):1749–1753, 2007.

[3] E Betzig, GH Patterson, R Sougrat, OW Lindwasser, S Olenych, JS Bonifacino, MW Davidson, J Lippincott-Schwartz, and Harald Hess. Imaging intracellular fluorescent proteins at nanometer resolution. *Science*, 313(5793):1642–1645, 2006.

[4] O Bieri, J Wirz, B Hellrung, M Schutkowski, M Drewello, and T Kiefhaber. The speed limit for protein folding measured by triplet-triplet energy transfer. *Proceedings of the National Academy of Sciences of the United States of America*, 96(17):9597–9601, 1999.

[5] R Brookmeyer, E Johnson, K Ziegler-Graham, and HM Arrighi. Forecasting the global burden of alzheimer's disease. *Alzheimer's & Dementia*, 3(3):186–191, 2007.

[6] M Bucciantini, E Giannoni, F Chiti, F Baroni, L Formigli, JS Zurdo, N Taddei, G Ramponi, CM Dobson, and M Stefani. Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. *Nature*, 416(6880):507–511, 2002.

[7] JH Cate, AR Gooding, E Podell, KH Zhou, BL Golden, CE Kundrot, TR Cech, and JA Doudna. Crystal structure of a group i ribozyme domain: Principles of rna packing. *Science*, 273(5282):1678–1685, 1996.

[8] K Chattopadhyay, EL Elson, and C Frieden. The kinetics of conformational fluctuations in an unfolded protein measured by fluorescence methods. *Proceedings of the National Academy of Sciences of the United States of America*, 102(7):2385–2389, 2005.

[9] K Chattopadhyay, S Saffarian, EL Elson, and C Frieden. Measurement of microsecond dynamic motion in the intestinal fatty acid binding protein

by using fluorescence correlation spectroscopy. *Proceedings of the National Academy of Sciences of the United States of America*, 99(22):14171–14176, 2002.

[10] HM Chen, E Rhoades, JS Butler, SN Loh, and WW Webb. Dynamics of equilibrium structural fluctuations of apomyoglobin measured by fluorescence correlation spectroscopy. *Proceedings of the National Academy of Sciences of the United States of America*, 104(25):10459–10464, 2007.

[11] DM Cheng, MM Dhingra, and RH Sarma. Spatial configuration of deoxyribotrinucleoside diphosphates in aqueous solution. *Nucleic Acids Research*, 5(11):4399–4416, 1978.

[12] F Chiti and CM Dobson. Protein misfolding, functional amyloid, and human disease. *Annual Reviews of Biochemistry*, 75:333–366, 2006.

[13] AE Cohen and WE Moerner. Method for trapping and manipulating nanoscale objects in solution. *Applied Physics Letters*, 86(9), 2005.

[14] PE Cole, DM Crothers, and SK Yang. Conformational changes of transfer ribonucleic acid - equilibrium phase diagrams. *Biochemistry*, 11(23):4358–&, 1972.

[15] MN Dessinges, B Maier, Y Zhang, M Peliti, D Bensimon, and V Croquette. Stretching single stranded dna, a model polyelectrolyte. *Physical Review Letters*, 89(24), 2002.

[16] S Doose, H Barsch, and M Sauer. Polymer properties of polythymine as revealed by translational diffusion. *Biophysical Journal*, 93(4):1224–1234, 2007.

[17] S Doose, H Neuweiler, and M Sauer. Fluorescence quenching by photoinduced electron transfer: A reporter for conformational dynamics of macromolecules. *Chemphyschem*, 10:1389–1398, 2009.

[18] DE Draper. A guide to ions and rna structure. *RNA*, 10(3):335, 2004.

[19] DE Draper. Rna folding: Thermodynamic and molecular descriptions of the roles of ions. *Biophysical Journal*, 95(12):5489–5495, 2008.

[20] L Edman, U Mets, and R Rigler. Conformational transitions monitored for single molecules in solution. *Proceedings of the National Academy of Sciences of the United States of America*, 93(13):6710–6715, 1996.

[21] M Eigen and R Rigler. Sorting single molecules - application to diagnostics and evolutionary biotechnology. *Proceedings of the National Academy of Sciences of the United States of America*, 91(13):5740–5747, 1994.

[22] D Eliezer and PE Wright. Is apomyoglobin a molten globule? structural characterization by nmr. *Journal of Molecular Biology*, 263(4):531–538, 1996.

[23] D Eliezer, J Yao, HJ Dyson, and PE Wright. Structural and dynamic characterization of partially folded states of apomyoglobin and implications for protein folding. *Nature Structural Biology*, 5(2):148–155, 1998.

[24] EL Elson and D Magde. Fluorescence correlation spectroscopy .1. conceptual basis and theory. *Biopolymers*, 13(1):1–27, 1974.

[25] M Fandrich, MA Fletcher, and CM Dobson. Amyloid fibrils from muscle myoglobin - even an ordinary globular protein can assume a rogue guise if conditions are right. *Nature*, 410(6825):165–166, 2001.

[26] A Fire, SQ Xu, MK Montgomery, SA Kostas, SE Driver, and CC Mello. Potent and specific genetic interference by double-stranded rna in caenorhabditis elegans. *Nature*, 391(6669):806–811, 1998.

[27] PJ Flory. *Principles of Polymer Chemistry*. Cornell University Press, Ithaca, NY, 1953.

[28] A Gelbin, B Schneider, L Clowney, SH Hsieh, WK Olson, and HM Berman. Geometric parameters in nucleic acids: Sugar and phosphate constituents. *Journal of the American Chemical Society*, 118(3):519–529, 1996.

[29] M Gotz, S Hess, G Beste, A Skerra, and ME Michel-Beyerle. Ultrafast electron transfer in the complex between fluorescein and a cognate engineered lipocalin protein, a so-called anticalin. *Biochemistry*, 41:4156–4164, 2002.

[30] BY Ha and D Thirumalai. Electrostatic persistence length of a polyelectrolyte chain. *Macromolecules*, 28(2):577–581, 1995.

[31] JH Ha and SN Loh. Changes in side chain packing during apomyoglobin folding characterized by pulsed thiol-disulfide exchange. *Nature Structural Biology*, 5(8):730–737, 1998.

[32] J Hardy and DJ Selkoe. The amyloid hypothesis of alzheimer's disease:

Progress and problems on the road to therapeutics. *Science*, 297(5580):353–356, 2002.

[33] U Haupts, S Maiti, P Schwille, and WW Webb. Dynamics of fluorescence fluctuations in green fluorescent protein observed by fluorescence correlation spectroscopy. *Proceedings of the National Academy of Sciences of the United States of America*, 95(23):13573–13578, 1998.

[34] SW Hell and J Wichmann. Breaking the diffraction resolution limit by stimulated-emission - stimulated-emission-depletion fluorescence microscopy. *Optics Letters*, 19(11):780–782, 1994.

[35] ST Hess, TPK Girirajan, and MP Mason. Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. *Biophysical Journal*, 91(11):4258–4272, 2006.

[36] ST Hess and WW Webb. Focal volume optics and experimental artifacts in confocal fluorescence correlation spectroscopy. *Biophysical Journal*, 83(4):2300–2317, 2002.

[37] FM Hughson, D Barrick, and RL Baldwin. Probing the stability of a partly folded apomyoglobin intermediate by site-directed mutagenesis. *Biochemistry*, 30(17):4113–4118, 1991.

[38] FM Hughson, PE Wright, and RL Baldwin. Structural characterization of a partly folded apomyoglobin intermediate. *Science*, 249(4976):1544–1548, 1990.

[39] C Bustamante I Tinoco Jr. How rna folds. *Journal of Molecular Biology*, 293:104–109, 2005.

[40] J Jacob, B Krantz, RS Dothager, P Thiyagarajan, and TR Sosnick. Early collapse is not an obligate step in protein folding. *Journal of Molecular Biology*, 338(2):369–382, 2004.

[41] M Jamin and RL Baldwin. Two forms of the ph 4 folding intermediate of apomyoglobin. *Journal of Molecular Biolpgy*, 276(2):491–504, 1998.

[42] J Jung and A Van Orden. Folding and unfolding kinetics of dna hairpins in flowing solution by multiparameter fluorescence correlation spectroscopy. *Journal of Physical Chemistry B*, 109(8):3648–3657, 2005.

[43] J Jung and A Van Orden. A three-state mechanism for dna hairpin folding characterized by multiparameter fluorescence fluctuation spectroscopy. *Journal of the American Chemical Society*, 128(4):1240–1249, 2006.

[44] MS Kay and RL Baldwin. Packing interactions in the apomyglobin folding intermediate. *Nature Structural Biology*, 3(5):439–445, 1996.

[45] GL Kenausis, J Voros, DL Elbert, NP Huang, R Hofer, L Ruiz-Taylor, M Textor, JA Hubbell, and ND Spencer. Poly(l-lysine)-g-poly(ethylene glycol) layers on metal oxide surfaces: Attachment mechanism and effects of polymer architecture on resistance to protein adsorption. *Journal of Physical Chemistry B*, 104(14):3298–3309, 2000.

[46] HD Kim, GU Nienhaus, T Ha, JW Orr, JR Williamson, and S Chu. Mg2+-dependent conformational change of rna studied by fluorescence correlation and fret on immobilized single molecules. *Proceedings of the National Academy of Sciences of the United States of America*, 99(7):4284–4289, 2002.

[47] J Kim, S Doose, H Neuweiler, and M Sauer. The initial step of dna hairpin folding: a kinetic analysis using fluorescence correlation spectroscopy. *Nucleic Acids Research*, 34(9):2516–2527, 2006.

[48] VN Kim. Microrna biogenesis: Coordinated cropping and dicing. *Nature Reviews Moleecular Cell Biology*, 6(5):376–385, 2005.

[49] S Kirmizialtin and R Elber. Computational exploration of mobile ion distributions around rna duplex. *Journal of Physical Chemistry B*, 114(24):8207–8220, 2010.

[50] O Kratky and G Porod. Rontgenuntersuchung geloster fadenmolekule. *Recueil des travaux chimiques des pays-bas - Journal of the Royal Netherlands Chemical Society*, 68(12):1106–1122, 1949.

[51] F Krieger, B Fierz, O Bieri, M Drewello, and T Kiefhaber. Dynamics of unfolded polypeptide chains as model for the earliest steps in protein folding. *Journal of Molecular Biology*, 332(1):265–274, 2003.

[52] EV Kuzmenkina, CD Heyes, and GU Nienhaus. Single-molecule forster resonance energy transfer study of protein dynamics under denaturing conditions. *Proceedings of the National Academy of Sciences of the United States of America*, 102(43):15471–15476, 2005.

[53] JR Lakowicz. *Principles of Fluorescence Spectroscopy*. Springer, 3rd edition, 2006.

[54] LJ Lapidus, WA Eaton, and J Hofrichter. Measuring single-molecule nucleic acid dynamics in solution by two-color filtered ratiometric fluorescence correlation spectroscopy. *Proceedings of the National Academy of Sciences of the United States of America*, 97(13):7220–7225, 2000.

[55] LJ Lapidus, PJ Steinbach, WA Eaton, A Szabo, and J Hofrichter. Effects of chain stiffness on the dynamics of loop formation in polypeptides. *Journal of Physical Chemistry B*, 106(44):11628–11640, 2002.

[56] TA Laurence, XX Kong, M Jager, and S Weiss. Probing structural heterogeneities and fluctuations of nucleic acids and denatured proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 102(48):17348–17353, 2005.

[57] CH Lee, FS Ezra, NS Kondo, RH Sarma, and SS Danyluk. Conformational properties of dinucleoside monophosphates in solution - dipurines and dipyrimidines. *Biochemistry*, 15(16):3627–3638, 1976.

[58] HT Li, XJ Ren, LM Ying, S Balasubramanian, and D Klenerman. Measuring single-molecule nucleic acid dynamics in solution by two-color filtered ratiometric fluorescence correlation spectroscopy. *Proceedings of the National Academy of Sciences of the United States of America*, 101(40):14425–14430, 2004.

[59] MA Lietzow, M Jamin, HJ Dyson, and PE Wright. Mapping long-range contacts in a highly unfolded protein. *Journal of Molecular Biology*, 322(4):655–662, 2002.

[60] L Lin, RJ Pinker, K Forde, GD Rose, and NR Kallenbach. Molten globular characteristics of the native state of apomyoglobin. *Nature Structural Biology*, 1(7):447–452, 1994.

[61] D Magde, EL Elson, and WW Webb. Fluorescence correlation spectroscopy .2. experimental realization. *Biopolymers*, 13(1):29–61, 1974.

[62] D Magde, WW Webb, and EL Elson. Thermodynamic fluctuations in a reacting system - measurement by fluorescence correlation spectroscopy. *Physical Review Letters*, 29(11):705–708, 1972.

[63] M Mandal and RR Breaker. Gene regulation by riboswitches. *Nature Reviews Molecular Cell Biology*, 5(6):451–463, 2004.

[64] M Manghi and RR Netz. Variational theory for a single polyelectrolyte chain revisited. *The European Physical Journal E: Soft Matter and Biological Physics*, 14:67–77, 2004.

[65] M Margittai, J Widengren, E Schweinberger, GF Schroder, S Felekyan, E Haustein, M Konig, D Fasshauer, H Grubmuller, R Jahn, and CAM Seidel. Single-molecule fluorescence resonance energy transfer reveals a dynamic equilibrium between closed and open conformations of syntaxin 1. *Proceedings of the National Academy of Sciences of the United States of America*, 100(26):15516–15521, 2003.

[66] N Marme, JP Knemeyer, M Sauer, and J Wolfrum. Inter- and intramolecular fluorescence quenching of organic dyes by tryptophan. *Bioconjugate Chemistry*, 14:1133–1139, 2003.

[67] ER McCarney, JH Werner, SL Bernstein, I Ruczinski, DE Makarov, PM Goodwin, and KW Plaxco. Site-specific dimensions across a highly denatured protein; a single molecule study. *Journal of Molecular Biology*, 352(3):672–682, 2005.

[68] DB McIntosh and OA Saleh. Salt species-dependent electrostatic effects on ssdna elasticity. *Macromolecules*, 44(7):2328–2333, 2011.

[69] KA Merchant, RB Best, JM Louis, IV Gopich, and WA Eaton. Characterizing the unfolded states of proteins using single-molecule fret spectroscopy and molecular simulations. *Proceedings of the National Academy of Sciences of the United States of America*, 104(5):1528–1533, 2007.

[70] U Meseth, T Wohland, R Rigler, and H Vogel. Resolution of fluorescence correlation measurements. *Biophys J*, 76(3):1619–1631, 1999.

[71] JB Mills, E Vacano, and PJ Hagerman. Flexibility of single-stranded dna: Use of gapped duplex helices to determine the persistence lengths of poly(dt) and poly(da). *Journal of Molecular Biology*, 285(1):245–257, 1999.

[72] W Min, GB Luo, BJ Cherayil, SC Kou, and XS Xie. Observation of a power-law memory kernel for fluctuations within a single protein molecule. *Physical Review Letters*, 94(19):198302, 2005.

[73] MC Murphy, I Rasnik, W Cheng, TM Lohman, and TJ Ha. robing single-stranded dna conformational flexibility using fluorescence spectroscopy. *Biophysical Journal*, 86(4):2530–2537, 2004.

[74] LJW Murray, WB Arendall, DC Richardson, and JS Richardson. Rna backbone is rotameric. *Proceedings of the National Academy of Sciences of the United States of America*, 100(24):13904, 2003.

[75] D Nettels, IV Gopich, A Hoffmann, and B Schuler. Ultrafast dynamics of protein collapse from single-molecule photon statistics. *Proceedings of the National Academy of Sciences of the United States of America*, 104(8):2655–2660, 2007.

[76] RR Netz and H Orland. Field theory for charged fluids and colloids. *Europhysics Letters*, 45:726–732, 1999.

[77] H Neuweiler, S Doose, and M Sauer. A microscopic view of miniprotein folding: Enhanced folding efficiency through formation of an intermediate. *Proceedings of the National Academy of Sciences of the United States of America*, 102(46):16650–16655, 2005.

[78] T Odijk. Polyelectrolytes near the rod limit. *Journal of Polymer Science B Polymer Physics*, 15:477–483, 1977.

[79] SA Pabit, X Qiu, JS Lamb, L Li, SP Meisburger, and L Pollack. Both helix topology and counterion distribution contribute to the more effective charge screening in dsrna compared with dsdna. *Nucleic Acids Research*, 37:3887–3896, 2009.

[80] AG Palmer and NL Thompson. Theory of sample translation in fluorescence correlation spectroscopy. *Biophysical Journal*, 51(2):339–343, 1987.

[81] N Panchuk-Voloshina, RP Haugland, J Bishop-Stewart, MK Bhalgat, PJ Millard, F Mao, WY Leung, and RP Haugland. Alexa dyes, a series of new fluorescent dyes that yield exceptionally bright, photostable conjugates. *Journal of Histochemistry and Cytochemistry*, 47:1179–1188, 1999.

[82] FG Parak and GU Nienhaus. Myoglobin, a paradigm in the study of protein dynamics. *Chemphyschem*, 3(3):249–254, 2002.

[83] L Pauling and RB Corey. Configurations of polypeptide chains with favored orientations around single bonds - 2 new pleated sheets. *Proceedings*

of the National Academy of Sciences of the United States of America, 37(11):729–740, 1951.

[84] JS Pedersen and P Schurtenberger. Scattering functions of semiflexible polymers with and without excluded volume effects. *Macromolecules*, 29(23):7602–7612, 1996.

[85] N. O. Petersen and EL Elson. Measurements of diffusion and chemical-kinetics by fluorescence photobleaching recovery and fluorescence correlation spectroscopy. *Methods Enzymology*, 130:454–484, 1986.

[86] KW Plaxco, IS Millett, DJ Segel, S Doniach, and D Baker. Chain collapse can occur concomitantly with the rate-limiting step in protein folding. *Nature Structural Biology*, 6(6):554–556, 1999.

[87] WH Press, SA Teukolsky, WT Vetterling, and BP Flannery. *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. Cambridge University Press, 3rd edition, 2007.

[88] AM Pyle. Ribozymes - a distinct class of metalloenzymes. *Science*, 261(5122):709–714, 1993.

[89] H Qian and EL Elson. Fluorescence correlation spectroscopy with high-order and dual-color correlation to probe nonequilibrium steady states. *Proceedings of the National Academy of Sciences of the United States of America*, 101(9):2828–2833, 2004.

[90] X Qiu, K Andresen, LW Kwok, JS Lamb, HY Park, and L Pollack. Inter-dna attraction mediated by divalent counterions. *Physical Review Letters*, 99(3), 2007.

[91] D Rehm and A Weller. Kinetics of fluorescence quenching by electron and h-atom transfer. *Israel Journal of Chemistry*, 8:259–271, 1970.

[92] R Rigler and EL Elson, editors. *Fluorescence Correlation Spectroscopy*. Chemical Physics. Springer-Verlag, Berlin, 1st edition, 2001.

[93] C Rischel, LE Jorgensen, and Z Foldes-Papp. Microsecond structural fluctuations in denatured cytochrome c and the mechanism of rapid chain contraction. *Journal of Physics:Condensed Matter*, 15(18):S1725–S1735, 2003.

[94] C Rischel, P Thyberg, R Rigler, and FM Poulsen. Time-resolved fluorescence studies of the molten globule state of apomyoglobin. *Journal of Molecular Biology*, 257(4):877–885, 1996.

[95] MJ Rust, M Bates, and X Zhuang. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (storm). *Nature Methods*, 3(10):793–795, 2006.

[96] JC Schlatterer, LW Kwok, JS Lamb, HY Park, K Andresen, M Brenowitz, and L Pollack. Hinge stiffness is a barrier to rna folding. *Journal of Molecular Biology*, 379:271–281, 1999.

[97] Y Seol, GM Skinner, and K Visscher. Elastic properties of a single-stranded charged homopolymeric ribonucleotide. *Physical Review Letters*, 93(11), 2004.

[98] Y Seol, GM Skinner, K Visscher, A Buhot, and A Halperin. Stretching of homopolymeric rna reveals single-stranded helices and base-stacking. *Physical Review Letters*, 98(15), 2007.

[99] J Skolnick and M Fixman. Electrostatic persistence length of a wormlike polyelectrolyte. *Macromolecules*, 10(5):944–948, 1977.

[100] SB Smith, YJ Cui, and C Bustamante. Overstretching b-dna: The elastic response of individual double-stranded and single-stranded dna molecules. *Science*, 271:795–799, 1996.

[101] SB Smith, L Finzi, and C Bustamante. Direct mechanical measurements of the elasticity of single dna-molecules by using magnetic beads. *Science*, 258(5085):1122–1126, 1992.

[102] L Stryer and RP Haugland. Energy transfer - a spectroscopic rules. *Proceedings of the National Academy of Sciences of the United States of America*, 58(2):719, 1967.

[103] M Sunde, LC Serpell, M Bartlam, PE Fraser, MB Pepys, and CCF Blake. Common core structure of amyloid fibrils by synchrotron x-ray diffraction. *Journal OF Molecular Biology*, 273(3):729–739, 1997.

[104] A Szabo, K Schulten, and Z Schulten. First passage time approach to diffusion controlled reactions. *Journal of Chemical Physics*, 72(8):4350–4357, 1980.

[105] MW Tate, EF Eikenberry, SL Barna, ME Wall, JL Lowrance, and SM Gruner. A large-format high-resolution area x-ray detector based on fiberoptically bonded charge-coupled-devide (ccd). *Journal of Applied Crystallography*, 28(2):196–205, 1995.

[106] O Tcherkasskaya and VN Uversky. Denatured collapsed states in protein folding: Example of apomyoglobin. *Proteins: Structure, Function and Genetics*, 44(3):244–254, 2001.

[107] D Thirumalai and BY Ha. *Statistical mechanicis in semiflexible chains*. Academic Press, 1998.

[108] B Tinland, A Pluen, J Sturm, and G Weill. Persistence length of single-stranded dna. *Macromolecules*, 30(19):5763–5765, 1997.

[109] AJ Trexler and E Rhoades. alpha-synuclein binds large unilamellar vesicles as an extended helix. *Biochemistry*, 48(11):2304–2306, 2009.

[110] AC Vaiana, H Neuweiler, A Schulz, J Wolfrum, M Sauer, and JC Smith. Fluorescence quenching of dyes by tryptophan: Interactions at atomic detail from combination of experiment and computer simulation. *Journal of the American Chemical Society*, 125:14564–14572, 2003.

[111] MI Wallace, LM Ying, S Balasubramanian, and D Klenerman. Fret fluctuation spectroscopy: Exploring the conformational dynamics of a dna hairpin loop. *Journal of Physical Chemistry B*, 104(48):11551–11555, 2000.

[112] S Weisbuch, F Gerard, M Pasdeloup, J Cappadoro, Y Dupont, and M Jamin. Cooperative sub-millisecond folding kinetics of apomyoglobin ph 4 intermediate. *Biochemistry*, 44(18):7013–7023, 2005.

[113] JH Werner, R Joggerst, RB Dyer, and PM Goodwin. A two-dimensional view of the folding energy landscape of cytochrome c. *Proceedings of the National Academy of Sciences of the United States of America*, 103(30):11130–11135, 2006.

[114] J Widengren, U Mets, and R Rigler. Fluorescence correlation spectroscopy of triplet-states in solution - a theoretical and experimental-study. *Journal of Physical Chemistry*, 99(36):13368–13379, 1995.

[115] J Widengren, R Rigler, and U Mets. Triplet-state monitoring by fluorescence correlation spectroscopy. *Journal of Fluorescence*, 4(3):255–258, 1994.

[116] SA Woodson. Metal ions and rna folding: a highly charged topic with a dynamic future. *Current Opinion in Chemical Biology*, 9:104–109, 2005.

[117] J Yao, J Chung, D Eliezer, PE Wright, and HJ Dyson. Nmr structural and dynamic characterization of the acid-unfolded state of apomyoglobin provides insights into the early events in protein folding. *Biochemistry*, 40(12):3561–3571, 2001.

[118] PR Young and NR Kallenbach. Secondary structure in polyuridylic acid - non-classical hydrogen-bonding and the function of the ribose 2'-hydroxyl group. *Journal of Molecular Biology*, 126(3):467–479, 1978.

[119] DP Zhong and AH Zewail. Femtosecond dynamics of flavoproteins: Charge separation and recombination in riboflavine (vitamin b-2)-binding protein and in glucose oxidase enzyme. *Proceedings of the National Academy of Sciences of the United States of America*, 98:11867–11872, 2001.